INTRODUCTION TO ARTIFICIAL NEURAL NETWORK AND MACHINE LEARNING

SEMINAR REPORT

submitted by

MUJEEB REHMAN O

for the award of the degree

of

Master of Technology

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
GOVERNMENT ENGINEERING COLLEGE SREEKRISHNAPURAM

PALAKKAD

January 2012

CERTIFICATE

GOVERNMENT ENGINEERING COLLEGE

SREEKRISHNAPURAM

PALAKKAD - 678 633

MCS 10 106(P) Seminar I

SEMINAR REPORT

This is to certify that this seminar report entitled **INTRODUCTION TO ARTIFICIAL NEURAL NETWORK AND MACHINE LEARNING** *submitted by* **Mujeeb Rehman O** *to the Department of Computer Science and Engineering, Government Engineering College, Sreekrishnapuram, Palakkad - 678633, in partial fulfilment of the requirement for the award of M.Tech Degree in Computational Linguistics is a bonafide record of the work carried out by him.*

Place: Sreekrishnapuram	Professor
Date:	Dept. of Computer Science & Engineering

Acknowledgement

First and foremost I wish to express my wholehearted indebtedness to God Almighty for his gracious constant care and blessings showered over me for the successful completion of the seminar.

I am indebted to Dr. K. Najeeb, Head of Department of Computer Science and Engineering, Govt. Engineering College Sreekrishnapuram, for providing and availing me all the required facilities for undertaking the seminar in a systematic way.

I am extremely greatful to Dr.P. C. Reghuraj, Professor, Department of Computer Science and Engineering, Govt Engineering College Sreekrishnapuram, for his sincere guidance, inspiration and right direction throughout the seminar.

Gratitude is extended to all teaching and non teaching staffs of Department of Computer Science and Engineering, Govt. Engineering College Sreekrishnapuram for the sincere directions imparted and the cooperation in connection with the seminar.

I am also thankful to my parents for the support given in connection with the seminar. Gratitude may be extended to all well-wishers and my friends who supported me to complete the seminar in time.

Table of Contents

List of Figures

List of Abbreviations

ANN	Artificial Neural Network
RBF	Radial Basis Function
PNN	Probabilistic Neural Network
GRNN	General Regression Neural Networks

Abstract

Artificial neural network (ANN) is a computational model that is inspired by the structural and functional aspects of biological neural networks. In this seminar, a brief introduction to models of ANN, model such as the single layer feed-forward network, multilayer feed-forward network, feedback network, and recurrent network will be given. This is followed by a discussion on McCulloch and Pitts Model and different types of activation functions.

Learning in ANN is a process by which the free parameters of the network are adapted through a process of stimulation by the environment in which the network is embedded. The type of learning is determined by the manner in which the parameter changes take place. This seminar also deals with the basics of learning strategies such as error correction, supervised, and unsupervised learning. We conclude with a discussion on Hebbian learning and Radial Basis Function (RBF) network.

CHAPTER 1

Introduction of Artificial Neural Network

Artificial Neural Network (ANN) is a mathematical model or computational model that is inspired by the structure and functional aspects of biological neural networks. A neural network consists of an interconnected group of artificial neurons. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. Modern neural networks are non-linear statistical data modeling tools. They are usually used to model complex relationships between inputs and outputs or to find patterns in data.

Because neuroscience is still full of unanswered questions, and since there are many levels of abstraction and therefore many ways to take inspiration from the brain, there is no single formal definition of what an artificial neural network is. Generally, it involves a network of simple processing elements that exhibit complex global behavior determined by connections between processing elements and element parameters. While an artificial neural network does not have to be adaptive per se, its practical use comes with algorithms designed to alter the strength (weights) of the connections in the network to produce a desired signal flow.

There are lot of formal definitions for ANN are there. Each of this definitions is focused its each properties of ANN. One of formal definitions is given below

Definition 1.1 *A neural network is a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experimental knowledge and making of available for use. It resembles the brain in two aspect:*

1. *Knowledge is acquired by the network from its environment through a learning process.*
2. *Inter-neuron connection strength, known as synaptic weight, are used to store the acquired knowledge[1].*

1.1 Benifits of ANN

1 Nonlinearity

2 Input-Output Mapping

3 Adaptivity

4 Evidential Respons

5 Contextual Information

6 Fault Tolarance

7 VLSI Implementability

8 Uniformity of Analysis and Design

9 Neurobiological Analogy

CHAPTER 2

Biological Neural Network

A biological neural network describes a population of physically interconnected neurons or a group of disparate neurons whose inputs or signaling targets define a recognizable circuit. Communication between neurons often involves an electrochemical process. The interface through which they interact with surrounding neurons usually consists of several dendrites (input connections), which are connected via synapses to other neurons, and one axon (output connection). If the sum of the input signals surpasses a certain threshold, the neuron sends an action potential (AP) at the axon hillock and transmits this electrical signal along the axon.

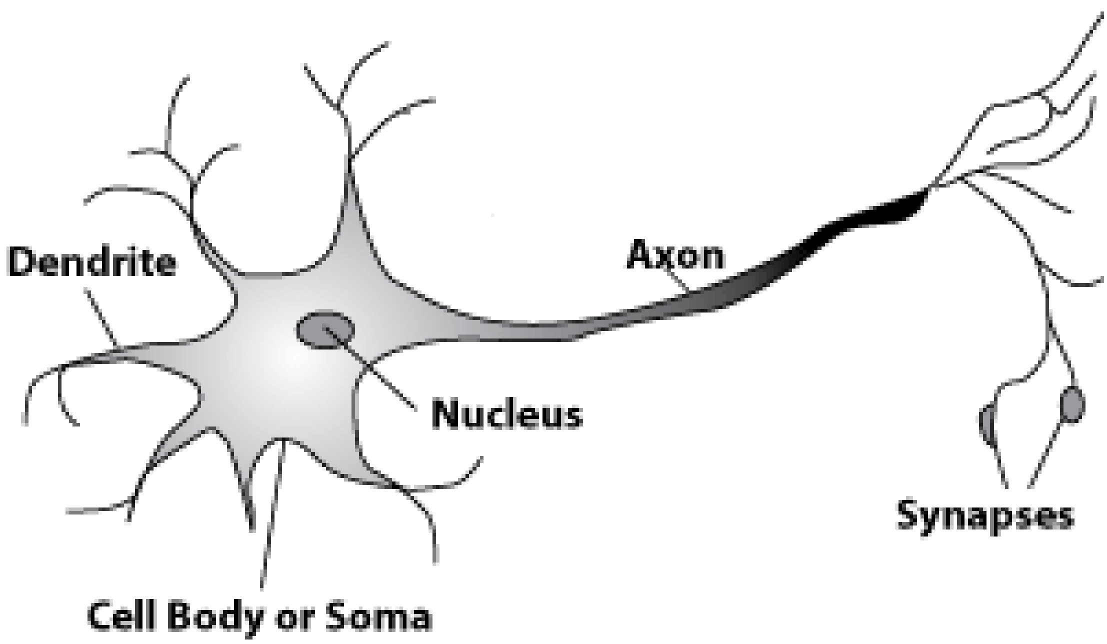

Figure 2.1: Biological Neuron

The figure shows the components of a neuron. Inter-connection of these neuron form biological neural network.Now we are going to discuss each components of neuron.

2.1 Dendrite

Dendrites are the branched projections of a neuron that act to conduct the electrochemical stimulation received from other neural cells to the cell body, or soma, of the neuron from which the dendrites project. Electrical stimulation is transmitted onto dendrites by upstream neurons via synapses which are located at various points throughout the dendritic arbor. Dendrites play a critical role in integrating these synaptic inputs and in determining the extent to which action potentials are produced by the neuron. Recent research has also found that dendrites can support action potentials and release neurotransmitters, a property that was originally believed to be specific to axons.

2.2 Soma

The soma is where the signals from the dendrites are joined and passed on. The soma and the nucleus do not play an active role in the transmission of the neural signal. Instead, these two structures serve to maintain the cell and keep the neuron functional.

The support structures of the cell include mitochondria, which provide energy for the cell, and the Golgi apparatus, which packages products created by the cell and secretes them outside the cell wall[2].

2.3 Axon

An axon (also known as a nerve fiber) is a long, slender projection of a nerve cell, or neuron, that typically conducts electrical impulses away from the neuron's cell body or soma.

An axon is one of two types of protoplasmic protrusions that extrude from the cell body of a neuron, the other type being dendrites. Axons are distinguished from dendrites by several features, including shape (dendrites often taper while axons usually maintain a constant radius), length (dendrites are restricted to a small region around the cell body while axons can be much longer), and function (dendrites usually receive signals while axons usually transmit them). All of these rules have exceptions, however.

Some types of neurons have no axon and transmit signals from their dendrites. No neuron ever has more than one axon; however in invertebrates such as insects or leeches the axon sometimes consists of several regions that function more or less independently of each other.Most axons branch, in some cases very profusely.

2.4 Synapse

In the nervous system, a synapse is a structure that permits a neuron to pass an electrical or chemical signal to another cell (neural or otherwise). The word "synapse" comes from "synaptein", which Sir Charles Scott Sherrington and colleagues coined from the Greek "syn-" ("together") and "haptein" ("to clasp").

Synapses are essential to neuronal function: neurons are cells that are specialized to pass signals to individual target cells, and synapses are the means by which they do so. At a synapse, the plasma membrane of the signal-passing neuron (the presynaptic neuron) comes into close apposition with the membrane of the target (postsynaptic) cell. Both the presynaptic and postsynaptic sites contain extensive arrays of molecular machinery that link the two membranes together and carry out the signaling process. In many synapses, the presynaptic part is located on an axon, but some presynaptic sites are located on a dendrite or soma.

There are two fundamentally different types of synapses:

- In a chemical synapse, the presynaptic neuron releases a chemical called a neurotransmitter that binds to receptors located in the postsynaptic cell, usually embedded in the plasma membrane. The neurotransmitter may initiate an electrical response or a secondary messenger pathway that may either excite or inhibit the postsynaptic neuron.

- In an electrical synapse, the presynaptic and postsynaptic cell membranes are connected by special channels called gap junctions that are capable of passing electrical current, causing voltage changes in the presynaptic cell to induce voltage changes in the postsynaptic cell. The main advantage of an electrical synapses is the rapid transfer of signals from one cell to the next[3].

CHAPTER 3

Basic Models of ANN

Neural networks are models of biological neural structures. The starting point for most neural networks is a model neuron. This neuron consists of multiple inputs and a single output. Each input is modified by a weight, which multiplies with the input value. The neuron will combine these weighted inputs and, with reference to a threshold value and activation function, use these to determine its output. This behavior follows closely our understanding of how real neurons work[4].

In ANN scenario we can model network on the basis of three entities. They are

- Synaptic interconnection
- Activation function
- Training or Learning

3.1 Synaptic interconnection

The Synapse is a bulb-like organ which is used to interconnect neurons.These types of model based on the interconnection of neurons(i.e., nodes). There exist five basic types of neuron connection architectures.
There are five types of network based on synaptic interconnection.

1. **Single layer feed-forward network:**

   The earliest kind of neural network is a single-layer feed forward network, which consists of a single layer of output nodes; the inputs are fed directly to the outputs via a series of weights. In this way it can be 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 if the value is above some threshold (typically 0) the neuron fires and takes the activated value (typically 1); otherwise it takes the deactivated value (typically -1). Neurons with this kind of activation function are also called Artificial neurons or linear threshold units. A similar neuron was described by Warren McCulloch and Walter Pitts in the 1940s. A example of single layered feed forward network is given below.

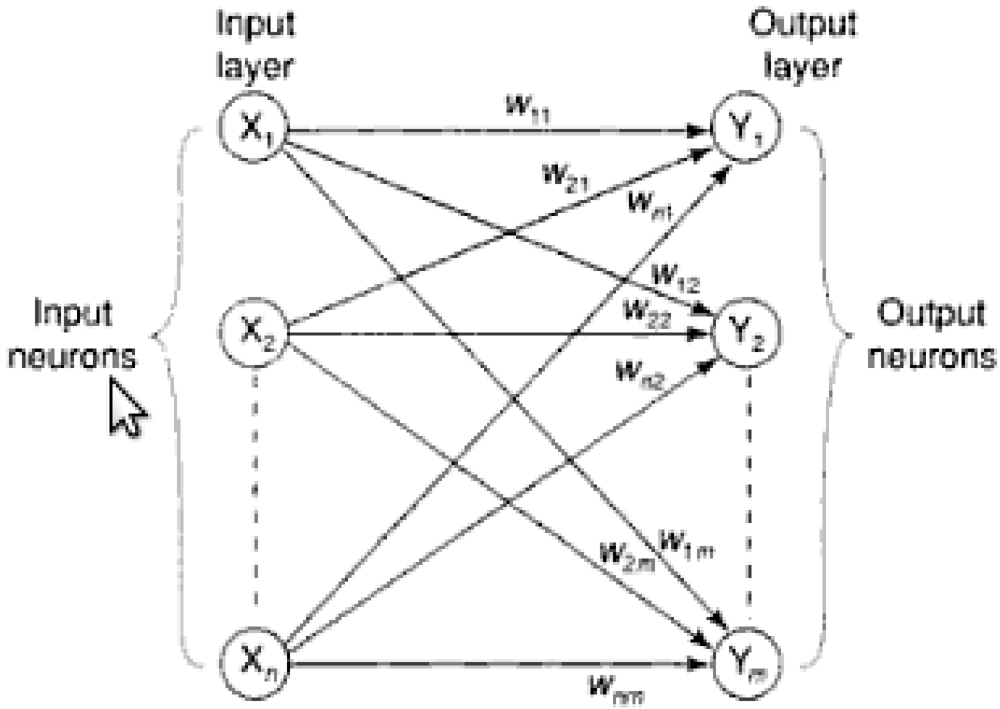

Figure 3.1: single layered feed-forward network

2. **Multilayer feed-forward network:**

   This class of networks consists of multiple layers of computational units, usually interconnected in a feed-forward way. Each neuron in one layer has directed connections to the neurons of the subsequent layer. There are input layer, hidden layers and an output layer consist in a multi layered feed-forward network .The multi layered feed-forward network is shown below.

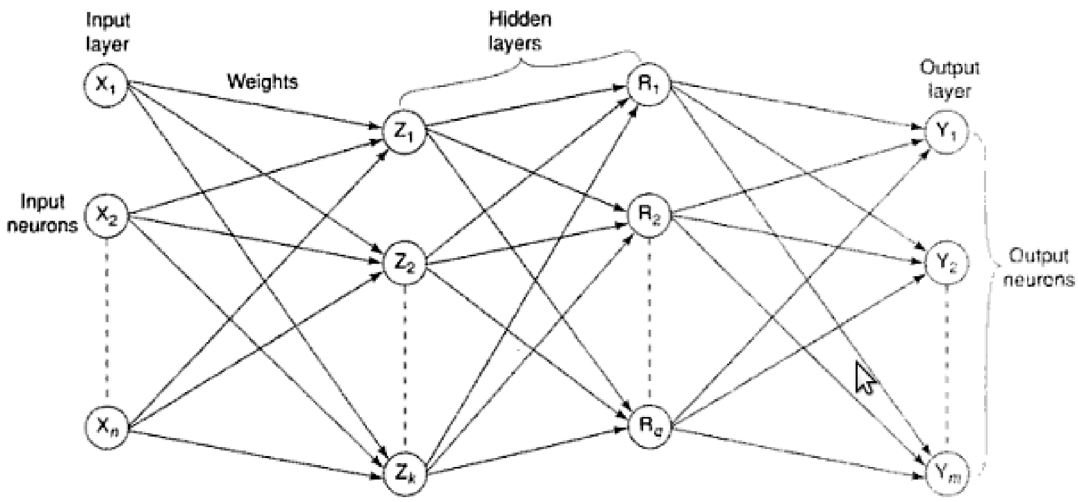

Figure 3.2: multi layered feed-forward network

3. **Single node with its own feedback:**

Feedback networks can have signals traveling in both directions by introducing loops in the network. Feedback networks are very powerful and can get extremely complicated. Feedback networks are dynamic; their 'state' is changing continuously until they reach an equilibrium point.They remain at the equilibrium point until the input changes and a new equilibrium needs to be found. Feedback architectures are also referred to as interactive or recurrent, although the latter term is often used to denote feedback connections in single-layer organizations[5].

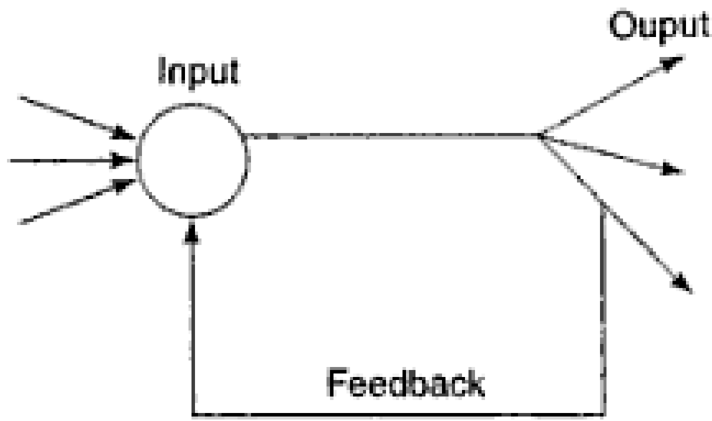

Figure 3.3: Single node with its own feedback

4. **Single-layer recurrent network:**

The output is fed back along with inputs of output layer.So this network can form a cyclic graph. This is a simple model in recurrent or feedback network. An example is given here.

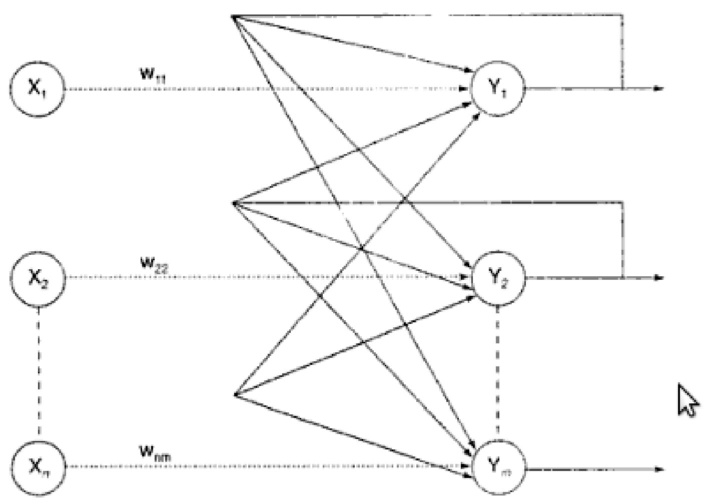

Figure 3.4: single-layer recurrent network

5. **Multilayer recurrent network**

The Multi layer recurrent network is consist of the properties of multi layer network as well as feedback network. A tipical example is given below.

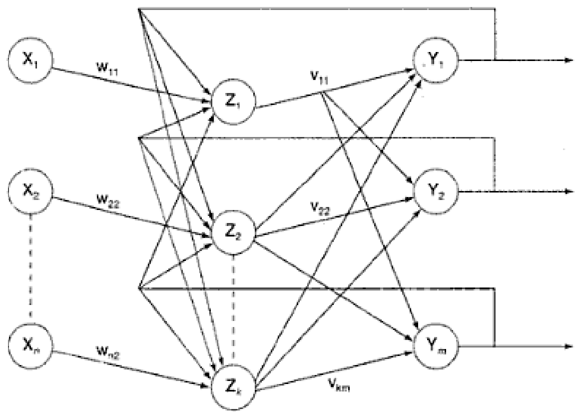

Figure 3.5: multi layer recurrent network

3.2 Activation functions

The activation function, denoted by $\varphi(v)$, defines the output of a neuron in terms of the induced local field v. There are five types of Activation functions. They are given below,

1. Identity function: it is a linear function defined as

$$f(x) = x \; \forall x \tag{3.1}$$

The output remains same as the input. The input layer uses the identity activation function.

2. Binary step function: This function can be defined as

$$f(x) = \{0 \; if \; x \; < \; \theta; \; 1 \; if \; x \; > \; \theta\} \tag{3.2}$$

Where θ the threshold variable. This function is mostly used in single-layer nets to convert the net input to an output that is binary (0 or 1).

3. Bipolar step function: This function can be defined as

$$f(x) = \{1\, if\ x\ <\ \theta;\ -1\, if\ x\ >\ \theta\} \tag{3.3}$$

Where θ the threshold variable. This function is also in single-layer nets to convert the net input to an output that is bipolar (+1 or -1).

4. Sigmoidal function: The sigmoidal function are widely used in back-propagation nets because of relationship between the value of the functions at a point and the value of derivative at that point which reduces the computational burden during training. Sigmoidal functions are of two types

   - *Binary Sigmoidal Function: This is also termed as logistic sigmoid function or unipolar sigmoid function. it can be defined as*

   $$f(x) = \frac{1}{1 + e^{-\lambda x}} \tag{3.4}$$

   Where λ is the steepness parameter. The derivative of this function is

   $$f'(x) = \lambda f(x)[1 - f(x)] \tag{3.5}$$

   Here the range of sigmoid function is from 0 to 1.

   - *Binary Sigmoidal Function: This function is defined as*

   $$f(x) = \frac{1 - e^{-\lambda x}}{1 + e^{-\lambda x}} \tag{3.6}$$

5. Ramp function:

$$f(x) = \{\ 1\, if\ x\ >\ 1\ ;\ x\, if\ 0\ \leqslant\ x\ \leqslant\ 1;\ 0\, if\ x\ <\ 0\ \} \tag{3.7}$$

CHAPTER 4

McCulloch-Pitts Model

The early model of an artificial neuron is introduced by Warren McCulloch and Walter Pitts in 1943. The McCulloch-Pitts neural model is also known as linear threshold gate. It is a neuron of a set of inputs $X_1, X_2, X_3, ..., X_m$ and one output y. The linear threshold gate simply classifies the set of inputs into two different classes. Thus the output y is binary. Such a function can be described mathematically using these equations:

$$s = \sum_{i=1}^{N} X_i W_i \tag{4.1}$$

There are two types of weights or give as input in McCulloch-Pitts neuron,*inhibitory* and *excitatory* .Inhibitory weights oppose the neuron to excite while excitatory helps.Here is a graphical representation of the McCulloch-Pitts model

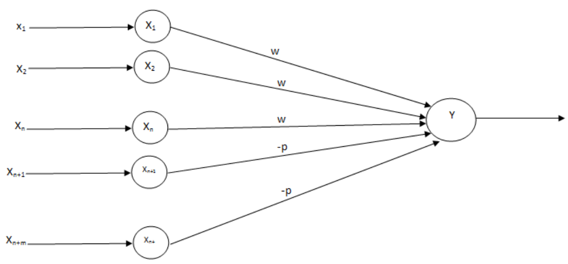

Figure 4.1: McCulloch-Pitts Architecture

Here x_1 to x_n possess excitatory weighted connection and x_{n+1} to x_{n+m} possess inhibitory weighted connection.The activation function used is binary step function.

If you want to operate the neurion in inhibition mode set θ in $\theta \; > \; nw \;\; p$ else set $kw \; \geq \; \theta \; > \; (k-1)w$

CHAPTER 5

Learning Process

Learning is a process by which the free parameters of a neural network are adapted through a process of stimulation by the environment in which the network is embedded. The type of learning is determined by the manner in which the parameter changes take place[1].

The learning process consist of three events. They are The Neural Network is stimulated, The Neural Network undergoes changes with respect to the stimulation and The Neural Network responds in a new way. Also we can broadly classify learning process as Parameter Learning and Structure learning.

There Describes five basic rule that control the learning process they are,

1. Error-Correction learning
2. Memory-Based learning
3. Hebbian learning
4. Competitive learning
5. Boltzmann learning

5.1 Hebbian learning

Hebb's postulate of learning is the oldest and most famous of all learning rule; it is named in honor of the neurophysiologist Hebb(1949).Hebb's theorem states that,

Theorem 5.1 *When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic changes takes place in one or both cells such that A's efficiency as one of the cells firing B, is increased[1].*

This statement is made in a neurobiological context. We may expand and rephrase it as a two-part rule.

1. If two neurons on either side of a synapse are activated *simultaneously*, then the strength of the synapse is increased.

2. If two neurons on either side of a synapse are activated *asynchronously*, then the synapse is weakened.

There are four mechanisms that affect Hebb's learning process.They are given below.

1. Time-dependent mechanism.
2. Local mechanism.
3. Interactive mechanism.
4. Conjunctional or correlational mechanism.

5.2 Types of Learning

The learning type can be classified as three. they are,

1. **Supervised learning**

   Supervised learning is the machine learning task of inferring a function from supervised (labeled) training data. The training data consist of a set of training examples. In supervised learning, each example is a pair consisting of an input object (typically a vector) and a desired output value (also called the supervisory signal). A supervised learning algorithm analyzes the training data and produces an inferred function, which is called a classifier (if the output is discrete, see classification) or a regression function (if the output is continuous, see regression). The inferred function should predict the correct output value for any valid input object. This requires the learning algorithm to generalize from the training data to unseen situations in a "reasonable" way.

2. **Unsupervised learning**

   Unsupervised learning studies how systems can learn to represent particular input patterns in a way that reflects the statistical structure of the overall collection of input patterns. By contrast with SUPERVISED LEARNING or REINFORCEMENT LEARNING, there are no explicit target outputs or environmental evaluations associated with each input; rather the unsupervised learner brings to bear prior biases as to what aspects of the structure of the input should be captured in the output.

3. **Reinforcement learning**

   Reinforcement learning (RL) is a kind of supervised learning in that some feedback from the environment is given. However the feedback signal is only evaluative, not instructive. Reinforcement learning is often called learning with a critic as opposed to learning with a teacher.

CHAPTER 6

Radial Basis Function

A Radial Basis Function (RBF) neural network has an input layer, a hidden layer and an output layer. The neurons in the hidden layer contain Gaussian transfer functions whose outputs are inversely proportional to the distance from the center of the neuron.

RBF networks are similar to K-Means clustering and PNN/GRNN networks. The main difference is that PNN/GRNN networks have one neuron for each point in the training file, whereas RBF networks have a variable number of neurons that is usually much less than the number of training points. For problems with small to medium size training sets, PNN/GRNN networks are usually more accurate than RBF networks, but PNN/GRNN networks are impractical for large training sets.

The RBF network consist of three layers.

1. The input layer made up of source nodes
2. The second layer made up of only hidden layer
3. The third layer consist of output nodes

6.1 Covers Theorem on the separability of patterns

When a RBF network is used to perform a complex pattern classification task, the problem is basically solved by transforming it into a high dimensional space in a nonlinear manner. The underlying justification is found in *Covers Theorem on the separability of patterns.*This may be stated as,

Theorem 6.1 *A complex pattern-classification problem cast in high-dimensional space nonlinearity is more likely to be linearly separable than in low-dimensional space[1].*

Consider a family of surface where each naturally divides an in put space into two regions. Let X denoted a set of N patterns $x_1, x_2, ...x_N$, each of which is assigned to one of two classes X_1 or X_2. This dichotomy (binary partition) of the points is said to be separable with respect to the family of surfaces if a surface exists in the family that separates the points in the class X_1 from those in the class X_2. For each pattern $x \in X$, define a vector made up of a set of real-valued functions [φ_i — i = 1, 2, ..., m_1], as shown by

$$\varphi(x) = [\varphi_1(x), \varphi_2(x), ...\varphi_m(x)]^T \tag{6.1}$$

Then, if x is a vector in an m_0-dimensional input space. Then $\varphi(x)$ maps to m_1-dimensional vector space. We can define a dichotomy $\{X_1\ ,X_2\}$ is said to be φ-separable if $\exists$ an m_1-dimensional vector w such that,

$$w^T\varphi(x) > 0 \tag{6.2}$$

if x $\in X_1$

$$w^T\varphi(x) < 0 \tag{6.3}$$

if x $\in X_2$

The hyperplane defined by the equation,

$$w^T\varphi(x) = 0 \tag{6.4}$$

6.2 An Example

An often quoted example which shows how the RBF network can handle a non-linearly separable function is the exclusive-or problem. In this problem there are four points (patterns), namely, (1,1), (0,1), (0,0) and (1,0), in a two dimensional input space. The requirement is to construct a pattern classifier that gives the binary input 0 in response to the input patterns (1,1) or (0,0), and the binary output 1 in response to the input pattern (0,1) or (1,0).

The solution has 2 inputs, 2 hidden units and 1 output. RBFs transform a nonlinearly separable problem into a linearly separable problem basically converting the input space into higher dimensional space. The pair of Gaussian hidden functions is given as follows:

$$\varphi_1 = e^{||xt_1||}, \quad t_1 = [1,1]^T \tag{6.5}$$

$$\varphi_2 = e^{||xt_2||}, \quad t_2 = [0,0]^T \tag{6.6}$$

where t1 and t2 are the centers of the gaussian functions. The results are summarized in Table 3.1 for the four different input patterns of interest. Correspondingly, the input patterns are mapped onto the 1-2 plane. There is no increase in the dimensionality of the hidden-unit space compared to the input space because the use of Gaussian hidden functions is satisfactory to transform the XOR problem into linearly separable one.

Input Pattern [x]	First Hidden Function [$\varphi_1(x)$]	Second Hidden Function [$\varphi_2(x)$]
(1, 1)	1	0.1353
(0, 1)	0.3678	0.3678
(0, 0)	0.1353	1
(1, 0)	0.36781	0.3678

CHAPTER 7

Conclusion

The concept of ANN is based on human brain mechanism .The ANN is used to solve more complex and non liner problems. Basic model of ANN is based on synaptic interconnection, learning rule and their activation function.Based on synaptic connection we classify it in five category.Also there are five types of activation functions.

Learning process means, adjusting the free parameters of a neural network to get a desired out put.There are five basic rules associated with learning process.And three types of learning process.

Bibliography

[**1**] Simon Haykin, *“Neural Network”*, Pearson Education

[**2**] *“http://psychology.about.com/od/biopsychology/ss/neuronanat_3.htm,”(visited on 11/02/2012)*

[**3**] Silverthorn, Dee Unglaub, William C. Ober, and Claire W. Garrison. *“Human Physiology, An Integrated Approach“*. Benjamin-Cummings Pub Co, 2007

[**4**] *”http://www.cheshireeng.com/Neuralyst/nnbg.htm”,(visited on 12/02/2012)*

[**5**]

“http://www.doc.ic.ac.uk/ñd/surprise_96/journal/vol4/cs11/report.html”,(visited on 12/02/2012)