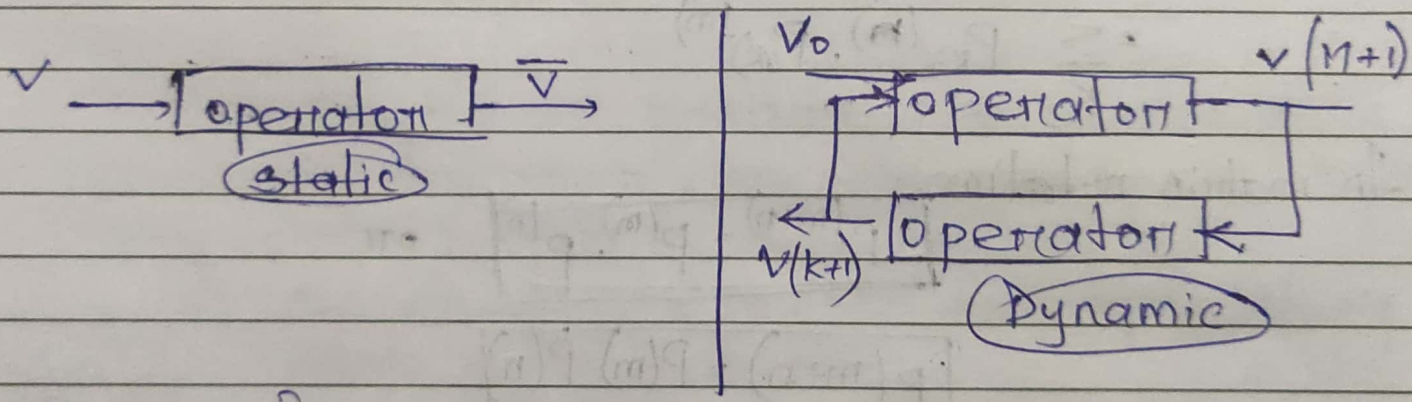
Associative memory: Associative memory networks can be seen as simplified model of human brain which can associate similar patterns.

⇒ It exhibits hebbian learning.
⇒ An associative memory is a store of associated patterns which are encoded in some form.

⇒ When storehouse is triggered with a pattern, the associated pair is recalled.

⇒ Associated memory is divided into two:

i) Static ⇒ recall output in one single feed forward pass.
ii) Dynamic ⇒ recalls through recurrent pass



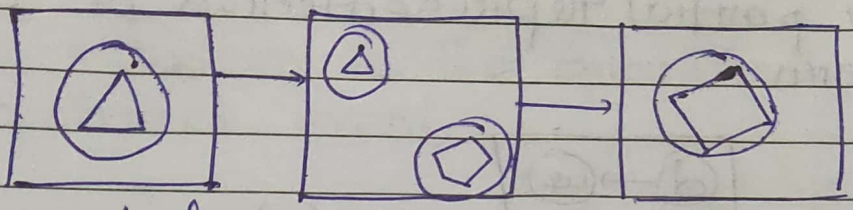
⇒ Types of associative memory:

i) Auto Associative: same input and output pair.
ii) Hetro Associative: associative patterns are different

Good Write

An associative memory is a content-addressable structure that maps a set of input patterns to a set of output patterns.

- Auto-associative memory: It recovers a previously stored pattern that most closely relates to the current pattern auto associative correlator:

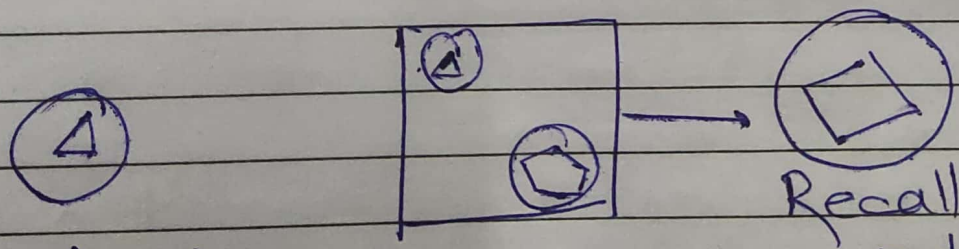


presented distorted pattern.

Recall of perfect pattern.

- Hetero-associative memory: The recovered pattern is generally different from the input pattern not only in type and format but also in content.

Also known as hetero-associative correlator.



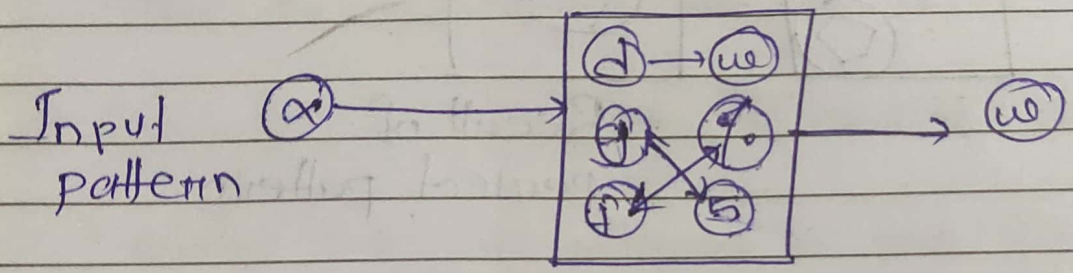
input pattern presented

Recall of associated pattern.

DATE: ___/___/____
PAGE: _____

Working of Associative Memory:

Associative memory is a depository of associated pattern which in some form. If the depository is triggered with a pattern the associated pattern pair appears at the ~~associated~~ output. The input would be an exact or partial representation of a stored pattern.



If the memory is produced with an input pattern, may say x, the associated pattern w is recovered automatically.

⇒ encoding or memorization
⇒ errors and noise
⇒ performance measure.

DATE: __/__/____
PAGE: _____

Mathematical Pattern:

Pattern term is used in context of neural networks to mean a set of activations across a pool of units (neurons).

ANN combines biological principles with advanced statistics to solve problems in domains such as pattern recognition.

Pattern → Pattern is the quantitative description of an object, event.

Patterns can be classified as: spatial or temporal pattern.

i) Spatial: Example are picture, weather, map, video, games.

ii) Temporal: Example are speech signal, ECG, etc. These pattern involves the ordered sequence of data appeared in time.

Pattern classification:

The main goal of pattern classification is to assign a physical object, event or phenomenon to one of the prespecified.

For example:

We classify various objects into different category like living, non-living thing etc.

Good Write

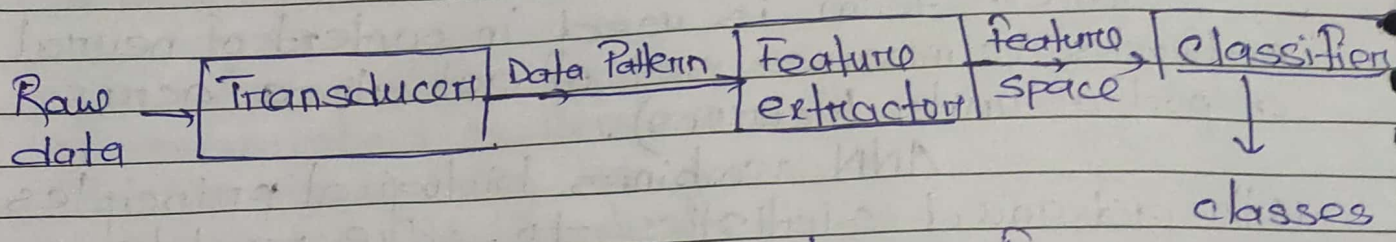


fig: Block diagram of classification

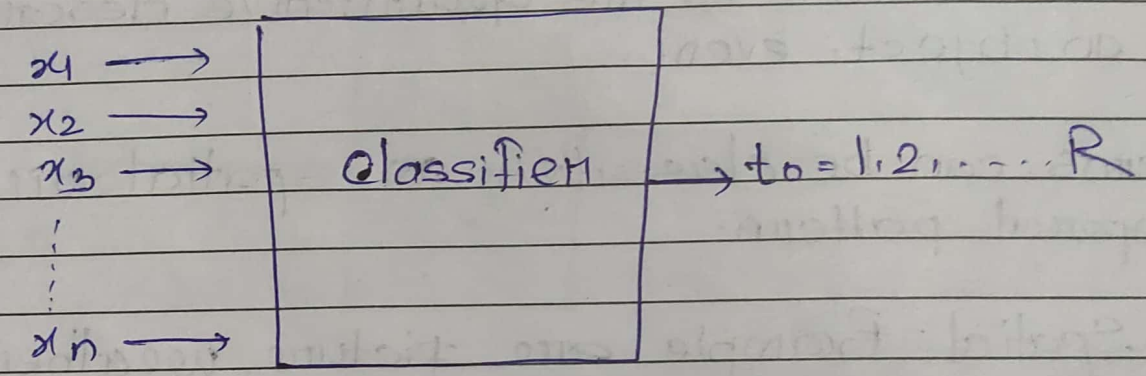


fig: General diagram of pattern classifier.

→ Transducer generally converts one form of energy to another form of energy.

Here, the input is raw data space $x = [x_1, x_2, \ldots x_n]$ and output is transducer is dia in pattern space that belong to certa category that is $t_o = 1, 2, \ldots R$.

Converted data is at the output can be compressed. Compressed data is called feature.

DATE: __/__/____
PAGE: _____

Hebbian learning: It was proposed by Donald O Hebb. It is one of the first and also easiest learning rule in the neural network used for pattern classification. It's a single layer NN having one input layer with n units and one output layer with one unit.

The rule works by updating the weights between neurons in the neural network for each training sample.

Hebbian learning Rule Algo:

1. Set all weights to zero, $w_i = 0$ for $i = 1$ to n and bias to 0.

2. For each input vector, s (input vector): t (target output pair), do step 3 to 5.

3. Set activation for input units with the input vector $x_i = s_i$ for $i = 1$ to n.

4. Set the corresponding output value to the output neuron, i.e. $y = t$.

5. Update weight and bias by applying Hebb rule for all $i = 1$ to n.

$$w_i(new) = w_i(old) + x_i y$$
$$b(new) = b(old) + y$$

DATE: __/__/____
PAGE: _____

Bidirectional Associative Memory:

BAM is a hetero-associative, content-addressable memory consist of 2 layers. It is similar to linear associator but connections are bidirectional that is,

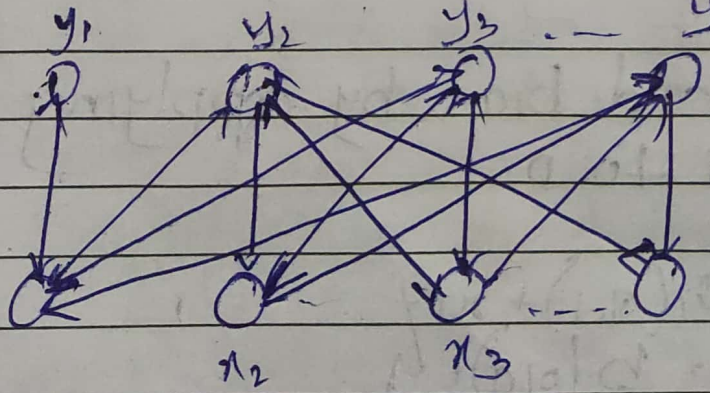
$$w_{ij} = w_{ji} \quad \text{for} \quad \begin{cases} i = 1, 2, \ldots m \\ j = 1, 2, \ldots n \end{cases}$$

- BAM allows forward and backward flow of information.

⇒ In BAM, the units in both layers serve as both input and output units depending on the direction of propagation.

⇒ Propagating the signals from X layer to Y layer makes the units in X layer to act as input units while units in Y layers act as output unit. Same is true for other direction because memory is bidirectional.

⇒ BAM model can perform both auto and hetero associative recall of stored info.



DATE: __/__/____
PAGE: _____

⇒ Like as linear associator and hopfield networks encoding in BAM can be carried out by using

$W_K = x_K^T y_K$ (to store a single associated pattern pair).

$W = \alpha \sum_{K=1}^{P} W_K$ (to store simultaneously several associated pattern pairs)

This process is generally called ~~storage on encoding~~ storage on encoding

After encoding network can be used for decoding. The decoding process is as follows:

1) Input layer can be applied, either on x layer or on the y layer depending on direction of propagation.

2) When input pattern is applied, network will propagate the input pattern to other layer allowing the units in other layer to compute their output value.

3) Pattern that was produced by other layer is propagated back to original layer and in original layer compute the output value.

4) This process is repeated until further propagation and computation do not result in a change in the associations of units.

Good Write

in both layer where final pattern pair is one of the stored associated pattern pair.

5) Final pattern pair that will be produced by the network depend on initial pattern pair and connection weight matrix.

Addition and Deletion of Pattern Pairs:

Given a set of pattern pairs (x_i, y_i) for $i = 1, 2, \ldots n$ and set of correlation Matrix M.

⇒ a new pair (x', y') can be added.
⇒ an existing pair (x_j, y_j) can be deleted from memory model.

~~Addition:~~

Addition: Add new pair (x', y') to existing correlation Matrix M, then new matrix M_{new} is given by,

$$M_{new} = x_1^T y_1 + x_2^T y_2 \ldots + x_n^T y_n + x^T y'$$

Deletion: Subtract the matrix corresponding to an existing pair (x_j, y_j) from matrix M, then correlation matrix is given by,

$$M_{new} = M - (x_j^T y_j)$$

DATE: __/__/____
PAGE: _____

- Discrete BAM: Network propagate an input pattern x to layer Y and Y layer compute the net input using:

Input is calculated. $y_i = \sum_{i=1}^{m} x_i w_{ij}$

O/P is calculated. $y_i(t+1) = \begin{cases} +1 & \text{if } y_i > 0 \\ y_i(t) & \text{if } y_i = 0 \\ -1 & \text{if } y_i < 0 \end{cases}$

Pattern Y produced on Y layer is then propagated back to X layer and X layer compute value:

Input, $x_j = \sum_{i=1}^{n} y_i w_{ij}$

O/P, $x_j(t+1) = \begin{cases} +1 & \text{if } x_i > 0 \\ x_i(t) & \text{if } x_i = 0 \\ -1 & \text{if } x_i < 0 \end{cases}$

- Continuous BAM: In this, the units use hyperboli tangent o/p function. The units in x-layers have an external input I_i while unit in Y-layer have an extra external input J_i.

For $i = 1, 2, \ldots m$ and $j = 1, 2, \ldots m$.

These extra input lead to a modification in the computation of net inputs.

$$x_i = \sum_{j=1}^{n} y_i w_{ij} + I_i$$

$$y_i = \sum_{j=1}^{m} x_i w_{ij} + J_i$$

DATE : __/__/____
PAGE : _____

Storage and Retrival Algorithm:

Storage algorithm is called encoding and retrieval algorithm is called decoding process.

Given,
P bipolar binary vector $\{s^{(1)}, s^{(2)}, \ldots s^{(p)}\}$

where, s^m is $n \times 1$ for $m = 1, 2, \ldots P$
Initializing vector v^0 is $n \times 1$.

Storage Algorithm $\Rightarrow$

Step 1: Weight matrix w is $(n \times n)$: $w \leftarrow 0$, $m \leftarrow 1$.

Step 2: Vector s^m is stored. So, the storage algo. for calculating the weight matrix is,

$$w \leftarrow s^m s^{m(t)}$$

or, $$w_{ij} = (1 - \delta_{ij}) \sum_{m=1}^{p} s_i(m)\, s_j(m)$$

Step 3: If $m < p$, then $m \leftarrow m+1$ and go to Step 2. Otherwise go to step 4.

Step 4: Storage is complete, output vector w.

Good Write

DATE: __/__/____
PAGE: _____

Retrival Algorithm:

Step 1: Cycle counter K is initialized $K \leftarrow 1$. Within the cycle counter i is initialized $i \leftarrow 1$ and initialized as $v \leftarrow v^0$.

Step 2: Integers 1, 2, ... n are arranged in an ordered random sequence if updation is not used then $\alpha_1 = 1, \alpha_2 = 2 \ldots \alpha_m = n$ or $\alpha_i = i$ at each update cycle.

Step 3: Neuron i is updated by computing

$v^{new}_{\alpha_i}$

$$net_{\alpha_i} = \sum_{j=1}^{n} W_{\alpha_i j} V_j$$

Step 4: If $i < n$ then $i \leftarrow i+1$ and go to Step 3, otherwise go to step 5.

Step 5: If $new\ x_j = V_{\alpha_i}$ for $i = 1, 2, \ldots n$.

DATE: __/__/____
PAGE: _____

Hopfield network: (Unsupervised learning)

⇒ It is a form of Recurrent artificial neural network invented by John Hopfield.
⇒ It is a type of ANN. (feedback)

⇒ Secure as content-addressable memory system with binary threshold units.

⇒ Types
- Discrete
- Continuous.

⇒ Hopfield network have a scalar value associated with each state of the network referred to as the "energy"

Discrete Hopfield network:

Discrete
- Bipolar
- Binary.

$$a_i \leftarrow \begin{cases} 1 & \text{if } \sum w_{ij} s_i > \theta_i \\ -1 & \text{otherwise} \end{cases}$$

$$a_i \leftarrow \begin{cases} 1 & \text{if } \sum w_{ij} s_i > \theta_i \\ 0 & \text{otherwise} \end{cases}$$

w_{ij} = strength of connection weight
s_j = state of unit j
θ_i = threshold of unit i

$$E = -\frac{1}{2} \sum_{y} w_j S_i S_j + \sum_{i} \theta_i S_i$$

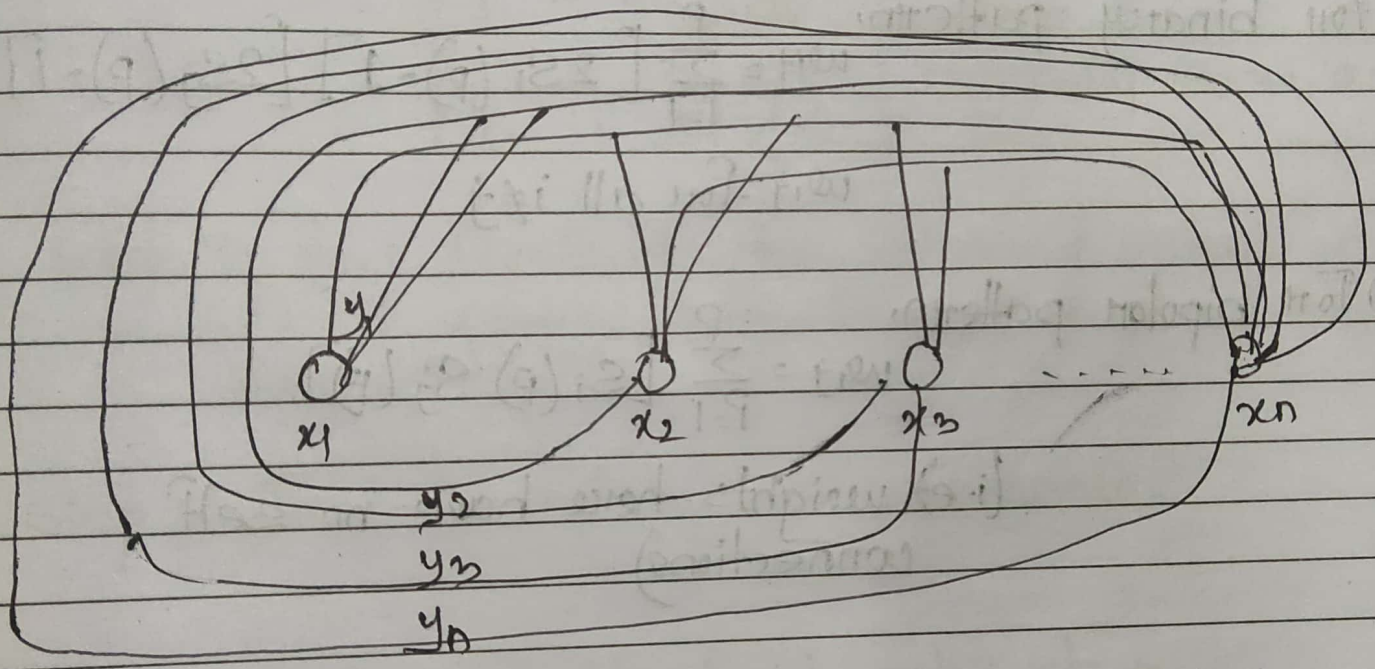
Good Write

DATE: __/__/____
PAGE: _____

Structure and Architecture.

⇒ Each neuron has an inverting and non-inverting output.

⇒ Being fully connected the output of each neuron is an output to all the neurons but not self.



Here, $x_1, x_2 \dots x_n$ → Input to the n given neuron.

$y_1, y_2, y_3 \dots y_n$ → output obtained from the n given neurons.

w_{ij} → weight associated with connection between the i^{th}

DATE : __/__/____
PAGE : _____

Training Algorithm:

For storing a set of input pattern $s(p)$ [$p = 1$ to P] where,

$S(p) = S_1(p) \ldots S_i(p) \ldots S_n(p)$, the weight matrix given by

For binary pattern,

$$w_{ij} = \sum_{p=1}^{P} [2S_i(p) - 1][2S_j(p) - 1]$$

w_{ij} for all $i \neq j$

For bipolar pattern,

$$w_{ij} = \sum_{p=1}^{P} [S_i(p) S_j(p)]$$

(i.e) weights here have no self connections)

Training Algorithm

step 1 = Initialize weights (w_{ij}) to store patterns (using training algo).

step 2 = For each input vector y_i, perform steps 3-8.

step 3 = Make initial activations of the network equal to the external input vector.

$y_i = x_i$ (for $i = 1$ to n).

step 4 = For each vector y_i, perform step 5-8

Good Write

DATE: __/__/____
PAGE: _____

step 5. Calculate the total input of the network y_{in} using the equation,

$$y_{in} = x_i + \sum_j [y_i w_{ji}]$$

step 6. Apply activation over the total input to calculate the output as per the equation

$$y_i = \begin{cases} 1 & \text{if } y_{in} > \theta_i \\ y_i & \text{if } y_{in} = \theta_i \\ 0 & \text{if } y_{in} < \theta_i \end{cases}$$

where θ is normally taken as 0.

step 7: Now, feedback the obtained output y_i to all other units. Thus, activation vectors are updated.

step 8: Test the network convergence.

Continuous Hopfield network ⇒

Here, the time parameter is treated as a continuous variable. So, instead of getting binary/bipolar outputs, we can obtain values that lie between 0 and 1.

It can be used to solve constrained optimization and associative memory problems.

$$v_i = g(u_i)$$

where, v_i = output from the continuous hopfield
u_i = internal activity of a node in continous

Good Write