WARSAW UNIVERSITY OF TECHNOLOGY
FACULTY OF MATHEMATICS
AND INFORMATION SCIENCE

Neural Networks

Lecture 9

Hamming Model

Hamming Model

The model with binary inputs and weights fixed during the preparatory phase.

The Hamming Model is a optimal classifier. The systems calculates the similarity (the Hamming distance) between the input signal and each pattern stored in the network. Next, the most similar stored pattern is selected.

Hamming Model

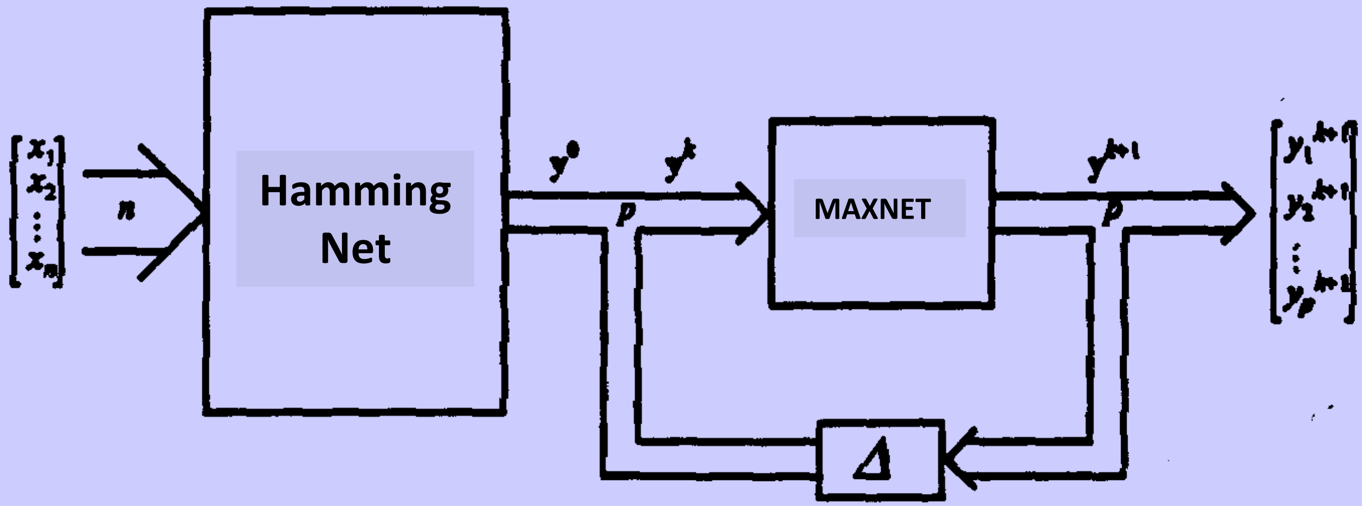
2 independent layers:

- the first layer – calculating the similarity
- the second layer – blocking all signals except the biggest one.

The weights and thresholds in the 1 layer are selected to assure that the s^{th} element input signal will be equal to $\boldsymbol{N - H^{input,s}}$ where $\boldsymbol{N}$ is the number of bits in the input signal (and of course in the stored patterns)), $H^{input,s}$ is the Hamming distance between the input signal and s^{th} stored pattern.

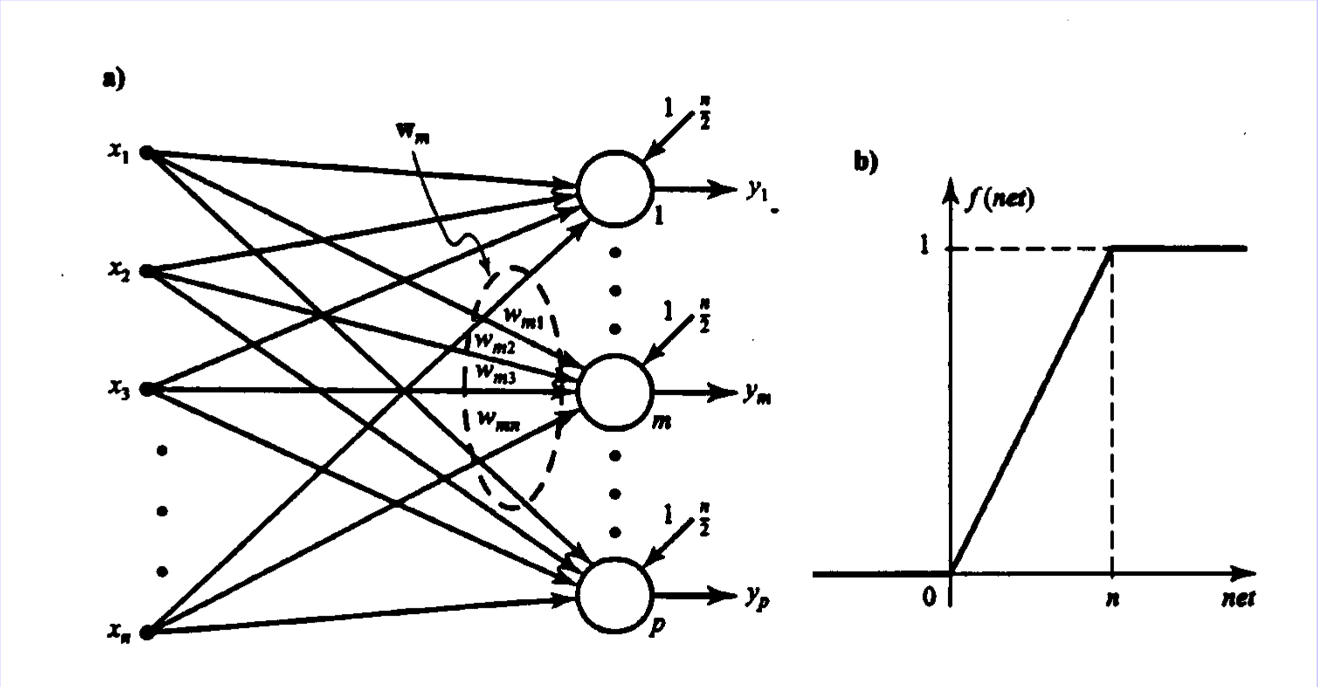
Hamming Model

Block diagram



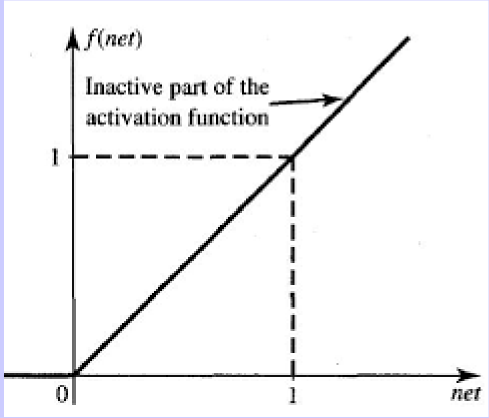
Hamming Model

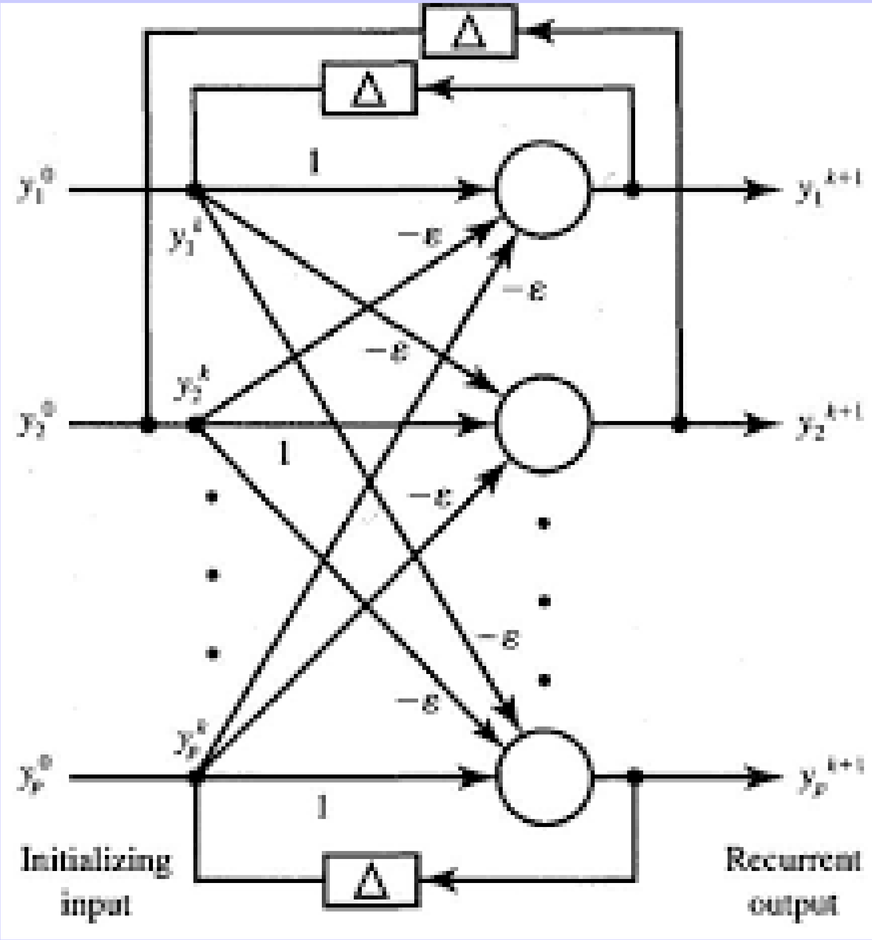
Hammings' network



Hamming Model

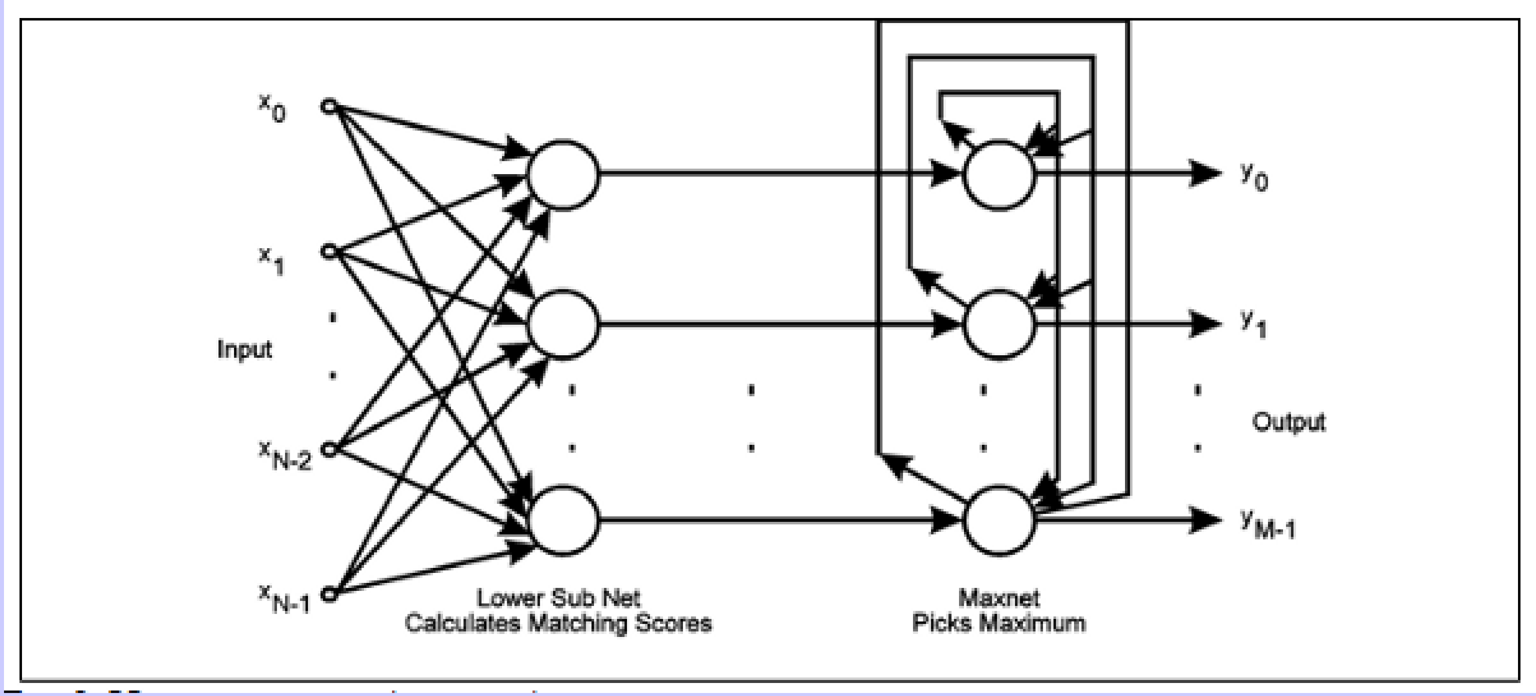
MAXNET





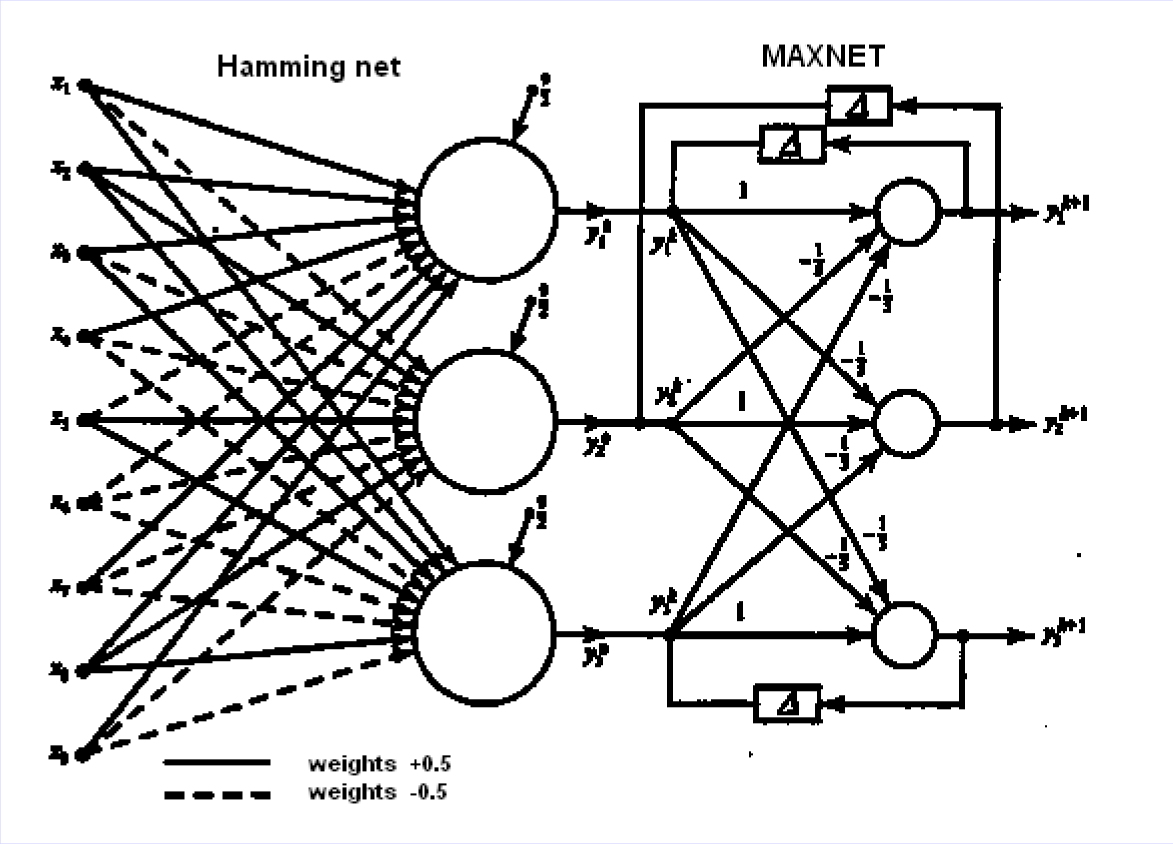
Hamming Model

Two layer Hamming network



Hamming Model

Two layer Hamming network (example)



Hamming Model

Output signals from Hamming's net are equal to: 1,2, ..., *N*. The greater value of output signal means that input signal ***X*** is more similar to the stored pattern *s*.

MAXNET, with internal connections based on the ***lateral inhibition rule*** has to select the greatest signal suppressing to zero the other signals.

Hamming Model

The net is able to store *p*, N-dimensional patterns $\mathbf{s}^{(m)}$.

Each element in the I layer is „*responsible*” for the one stored pattern. The incoming weights to that *m*-th element

$$\mathbf{w}_m = [w_{m1}, w_{m2}, \ldots, w_{mN}]$$

connects input elements with that element.

The classifierer has *p* class, *p* elements and *p* outputs.

Hamming Model

The nonlinear characteristic of an element produce at the output the signal is equal to 1 if and only if the input signal is identical with the stored pattern.

The incoming weights of the element (*m*), that one where the *m*-ty pattern is stored, are equal

$$\mathbf{w}_m = \mathbf{s}^{(m)}$$

The input signals of network elements are

$$\mathbf{X}^T\mathbf{s}^{(1)}, \mathbf{X}^T\mathbf{s}^{(2)}, \ldots , \mathbf{X}^T\mathbf{s}^{(m)}, \ldots , \mathbf{X}^T\mathbf{s}^{(p)}$$

Hamming Model

If the input signal $\mathbf{X} = \mathbf{s}^{(m)}$, the only one weighted input is equal to **N,** and the rest belongs to the (-N;+N) (the input signals x_i are equal to -1 or +1).

The inner (scalar) products $\mathbf{X}^T\mathbf{s}^{(m)}$ are used to calculate the similarities between the input signal and stored patterns.

The inner product $\mathbf{X}^T\mathbf{s}^{(m)}$ can be written as:

the number of positions (bits) where they agree – minus the number of positions where they disagree.

Hamming Model

The number of positions they disagree – it is Hammings' distance

$$H(\mathbf{X},\mathbf{s}^{(m)})$$

so, the number where they agree

$$N - H(\mathbf{X},\mathbf{s}^{(m)})$$

hence $\mathbf{X}^T\mathbf{s}^{(m)} = \left\{N - H(\mathbf{X},\mathbf{s}^{(m)})\right\} - H(\mathbf{X},\mathbf{s}^{(m)})$,

$$\boxed{\mathbf{X}^T\mathbf{s}^{(m)}/2 = N/2 - H(\mathbf{X},\mathbf{s}^{(m)})}$$

Hamming Model

The weight matrix of Hammings' net **W** (connections with the I layer)

$$\boldsymbol{W} = \frac{1}{2}\begin{bmatrix} \boldsymbol{s}_1^{(1)} & \boldsymbol{s}_2^{(1)} & \circ & \circ & \boldsymbol{s}_{\boldsymbol{N}}^{(1)} \\ \boldsymbol{s}_1^{(2)} & \boldsymbol{s}_2^{(2)} & \circ & \circ & \boldsymbol{s}_{\boldsymbol{N}}^{(2)} \\ \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ \\ \boldsymbol{s}_1^{(\boldsymbol{p})} & \boldsymbol{s}_2^{(\boldsymbol{p})} & \circ & \circ & \boldsymbol{s}_{\boldsymbol{N}}^{(\boldsymbol{p})} \end{bmatrix}$$

The input signal **X** produce at the input of each element signal

$$\frac{1}{2}\boldsymbol{X}\boldsymbol{s}^{(\boldsymbol{m})}$$

Hamming Model

Plus the additional constant bias signal of N/2,

$$E^H = \frac{1}{2} X s^{(m)} + \frac{N}{2} = N - H(X, s^{(m)})$$

The nonlinear characteristic

$$f(E^H) = \frac{1}{N} E^H$$

produce at the output signal of value <0; 1>. The element of a greater output signal indicates the class (the number of a class) where the input signal **X** has the smallest Hamming distance. The best matching

H = 0 i ***f(E^H) =1***

Hamming Model

The iterative procedure of MAXNET have to suppress the rest (smaller) output signals

$$\boldsymbol{W}_{MAXNET} = \begin{bmatrix} 1 & -\varepsilon & \circ & \circ & -\varepsilon \\ -\varepsilon & 1 & \circ & \circ & -\varepsilon \\ \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ \\ -\varepsilon & -\varepsilon & \circ & \circ & 1 \end{bmatrix} = \boldsymbol{W}_M$$

where $0 < \varepsilon < 1/p$

the coefficient of lateral inhibition

Hamming Model

The recursive procedure

$$\boldsymbol{y}(\boldsymbol{t}+1)=\Psi\left[\boldsymbol{W}_{\boldsymbol{MAXNET}}\cdot\boldsymbol{y}(\boldsymbol{t})\right]$$

where

$$\Psi[\mathrm{a}]=\begin{cases}0 \text{ if } \mathrm{a}<0\\ \mathrm{a} \text{ if } \mathrm{a}\geq 0\end{cases}$$

Hamming Model

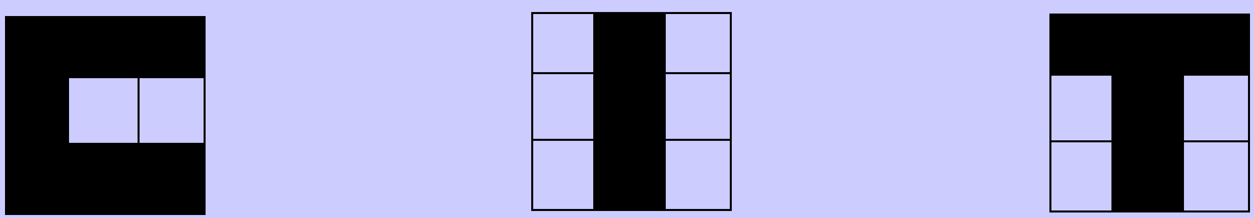
Example

$$\boldsymbol{s}^{(1)} = \begin{bmatrix} +1 & +1 & +1 & +1 & -1 & -1 & +1 & +1 & +1 \end{bmatrix}$$

$$\boldsymbol{s}^{(2)} = \begin{bmatrix} -1 & +1 & -1 & -1 & +1 & -1 & -1 & +1 & -1 \end{bmatrix}$$

$$\boldsymbol{s}^{(3)} = \begin{bmatrix} +1 & +1 & +1 & -1 & +1 & -1 & -1 & +1 & -1 \end{bmatrix}$$

these are patterns of C, I and T



Hamming Model

the weight vector

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{s}^{(1)} \\ \boldsymbol{s}^{(2)} \\ \boldsymbol{s}^{(3)} \end{bmatrix}$$

the input signals to the Hamming net elements

$$\boldsymbol{E}^{\boldsymbol{H}} = \frac{1}{2}\boldsymbol{W} \cdot \boldsymbol{X} + \begin{bmatrix} 9/2 \\ 9/2 \\ 9/2 \end{bmatrix}$$

Hamming Model

Lat us assume that the input signal

$$\boldsymbol{X} = \begin{bmatrix} +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 \end{bmatrix}$$

it has to be classifier to one of three classes ***C,I,T***

Let $\varepsilon = 0{,}2$ (< 1/3 = 1/p), hence

$$\boldsymbol{E}^{\boldsymbol{H}} = \begin{bmatrix} 7 \\ 3 \\ 5 \end{bmatrix} \qquad \boldsymbol{f}(\boldsymbol{E}^{\boldsymbol{H}}) = \begin{bmatrix} 7/9 \\ 3/9 \\ 5/9 \end{bmatrix}$$

$f(\boldsymbol{E}^H) = y(0)$ is the first input to the MAXNET

Hamming Model

Iterative procedure

$$\boldsymbol{y}(\boldsymbol{t}+1)=\Psi\left[\boldsymbol{E}^{M}\cdot\boldsymbol{y}(\boldsymbol{t})\right]$$

yields

$$\boldsymbol{W}_{M}\cdot\boldsymbol{y}(\boldsymbol{t})=\begin{bmatrix}1 & -0.2 & -0.2\\ -0.2 & 1 & -0.2\\ -0.2 & -0.2 & 1\end{bmatrix}\begin{bmatrix}\boldsymbol{y}_{1}(\boldsymbol{t})\\ \boldsymbol{y}_{2}(\boldsymbol{t})\\ \boldsymbol{y}_{3}(\boldsymbol{t})\end{bmatrix}$$

and next

$$\boldsymbol{E}^{M}(1)=\left\lfloor {}^{6}\!/_{10} \quad {}^{1}\!/_{15} \quad {}^{1}\!/_{3}\right\rfloor=\boldsymbol{y}(1)$$

Hamming Model

next

$$E^M(2) = \begin{bmatrix} \frac{13}{25} & \frac{-3}{25} & \frac{1}{5} \end{bmatrix} \Rightarrow$$

$$y(2) = \begin{bmatrix} \frac{13}{25} & 0 & \frac{1}{5} \end{bmatrix}$$

$$E^M(3) = \begin{bmatrix} \frac{12}{25} & \frac{-18}{125} & \frac{12}{125} \end{bmatrix} \Rightarrow$$

$$y(3) = \begin{bmatrix} \frac{12}{25} & 0 & \frac{12}{125} \end{bmatrix}$$

$$E^M(4) = \begin{bmatrix} \frac{576}{1250} & \frac{-144}{1250} & 0 \end{bmatrix} \Rightarrow$$

$$y(4) = \begin{bmatrix} \frac{576}{1250} & 0 & 0 \end{bmatrix}$$

Hamming Model

The last result stops the procedure – the stable state is achieved.

Conclussion:

The unknown input signal ***X*** had the smallest Hamming distance from the pattern $s^{(1)}$, so it belonged to the

class - **C**