ARTIFICIAL NEURAL NETWORK

Unit-2: Perceptron

Ms. Swetha R.
Department of Electronics and Communication Engineering
PES University

CONTENT:Part-1

1. *Perceptron*
    1. *Introduction-Linearly Separable*
    2. *Rosenblatt Algorithm with example*
    3. *Perceptron Convergence Theorem*
2. *Single Layer Perceptron*

    DrawBack: Xor Logic Gate
3. *Multilayer Perceptron*
    1. *Backpropogation Algorithm*
    2. *Example: XOR Logic Gate*

Perceptron Convergence Theorem:

If there is a weight vector W* such that

$$\varphi\left(W^* X(k)\right) = y(k)$$

then for any starting vector W, the perceptron learning rule will converge to weight vector W* that gives the correct response for all training patterns and it will do so in finite number of steps

Perceptron Convergence Theorem

Assumptions:

- Inputs to the perceptron originate from 2 linearly separable classes.
- W(0)=0
- Learning rate is 1 and it remains constant

Perceptron Convergence Theorem

Proof:

- Let
  - C_1 and C_2 be 2 different classes,
  - $H_1 \subset C_1$ and $H_2 \subset C_2$
- Define $X(n) = [+1, x_1(n), x_2(n), \cdots, x_m(n)]^T$

$$W(n) = [b(n), w_1(n), w_2(n), \cdots, w_m(n)]^T$$

- Require a 'W' such that,

$$w \in R^{m+1},$$

$$v(x) = W^T X \geq 0 \forall X \in H_1$$

$$v(x) = W^T X < 0 \forall X \in H_2$$

Perceptron Convergence Theorem

For $k = 0$

$$v(0) = W^T(0X(0) = 0$$

$$\varphi(v) = \begin{cases} 1 & v \geq 0 \\ 0 & v < 0 \end{cases}$$

if $X(0) \in H_1$, then $W(1) = W(0)$

otherwise $x(1) \in H_2$, then $W(2) = W(1) - X(1) = -X(1)$

To understand in easier way, i assume that the perceptron incorrectly classifies the vectors x(1),X(2)... Therefore,

$$W^T(k)X(k) < 0 \ for \quad k = 1,2,...and$$

$$the \quad input \quad vector \quad X(k) \in H_1$$

Perceptron Convergence Theorem

For $k = 1$

$$v(1) = W^T(1)X(1) < 0$$

since $X(1) \in H_1$ $\quad W(2) = W(1) + X(1) = X(1)$

For $k = 2$

$$v(2) = W^T(2)X(2) < 0$$

$X(2) \in H_1$ and

$$W(3) = W(2) + X(2) = X(1) + X(2)$$

At the kth stage we obtain

$$W(k+1) = X(1) + X(2) + \ldots. + X(k)$$

Since the Classes C1 and C2 are assumed to be linearly separable, there exists a solution $W_{0.}$

Then for a fixed solution W_0, we may define a positive

$$\alpha = \min_{X(i) \in H_1} W_0^{\,T} X(i)$$

$$W(k+1) = X(1) + X(2) + \ldots. + X(k)$$

Multiply W_0 on both the side of the equation

$$W_0^T W(k+1) = W_0^T X(1) + W_0^T X(2) + \ldots. + W_0^T X(k)$$

$$W_0^T W(k+1) \geq k\alpha \qquad \text{.......................(1)}$$

Let $x, y \in R^m$, then $\left|x^T y\right| \le \|x\|^2 \|y\|^2$

- The above inequality is referred as Cauchy-Schwarz inequality

Therefore, $\left|W_0^T W(k+1)\right| \le \|W_0\|^2 \|W(k+1)\|^2$

$$\|W(k+1)\|^2 \ge \frac{k^2 \alpha^2}{\|W_0\|^2} \qquad \cdots\cdots(a)$$

Under the initial assumption

$$W(k+1) = W(k) + X(k) \qquad for \quad k = 1,2\ldots$$

By taking the squared Euclidean norm of both sides of above equation

$$\|W(k+1)\|^2 = \|W(k)\|^2 + \|X(k)\|^2 + 2W^T(k)X(k)$$

$$\|W(k+1)\|^2 \le \|W(k)\|^2 + \|X(k)\|^2$$

$$\|W(k+1)\|^2 - \|W(k)\|^2 \le \|X(k)\|^2$$

$$\|W(k+1)\|^2 \le \sum_{X(i)\in C_1} \|X(k)\|^2$$

Let

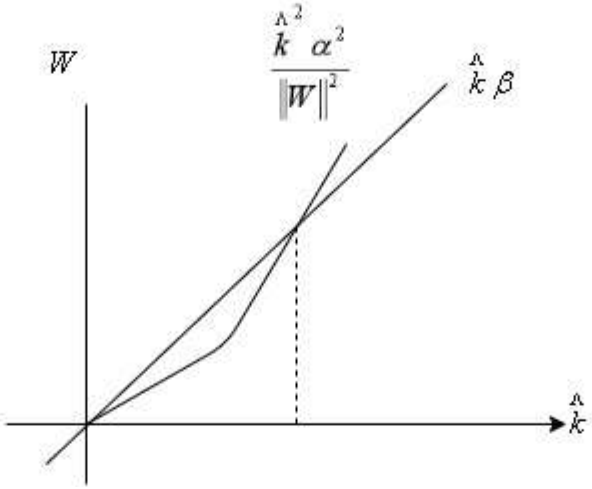
$$\beta = \max_{X(k)\in C_1} \|X(k)\|^2$$

$$\Rightarrow \|W(k+1)\|^2 \le k\beta \qquad \cdots\cdots(b)$$

Compare (a) and (b)

$$\frac{k^2\alpha^2}{\|W_0\|^2} \le \|W(k+1)\|^2 \le k\beta$$

$$k_{\max} = \frac{\beta \|W\|^2}{\alpha^2}$$



Artificial Neural Network-Percepton

Single Layer Perceptron

- Now lets consider 2 input XOR logic gate.
- Is it possible to design the gate using Single layer perceptron?

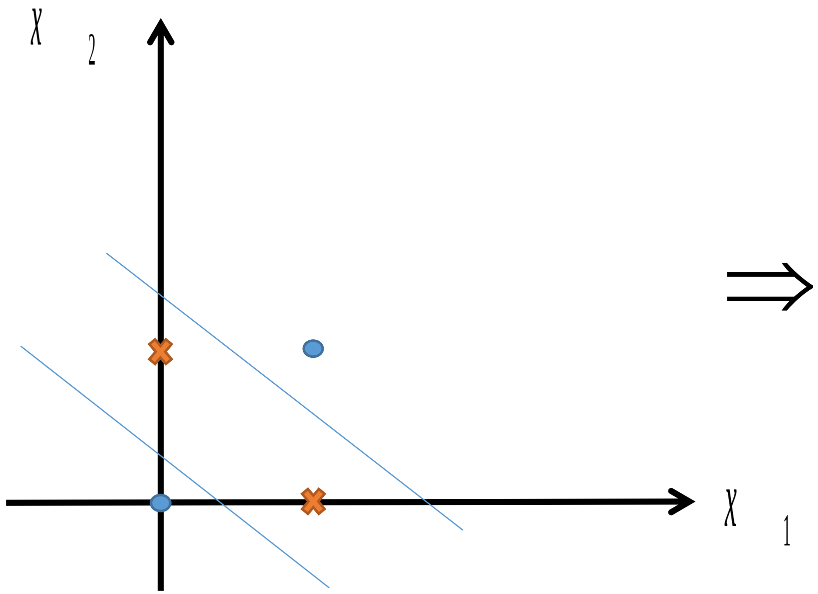
x	y	z
0	0	0
1	0	1
0	1	1
1	1	0

C_1

C_2

Artificial Neural Network-Perceptron

Single Layer Perceptron



$\Rightarrow$

- Therefore, Single-layer Perceptron cannot be used in this case.
- This problem can be solved using Multi-layer Perceptron

THANK YOU

Ms. Swetha R.
Department of Electronics and Communication Engineering

swethar@pes.edu

+91 80 2672 1983 Extn 753