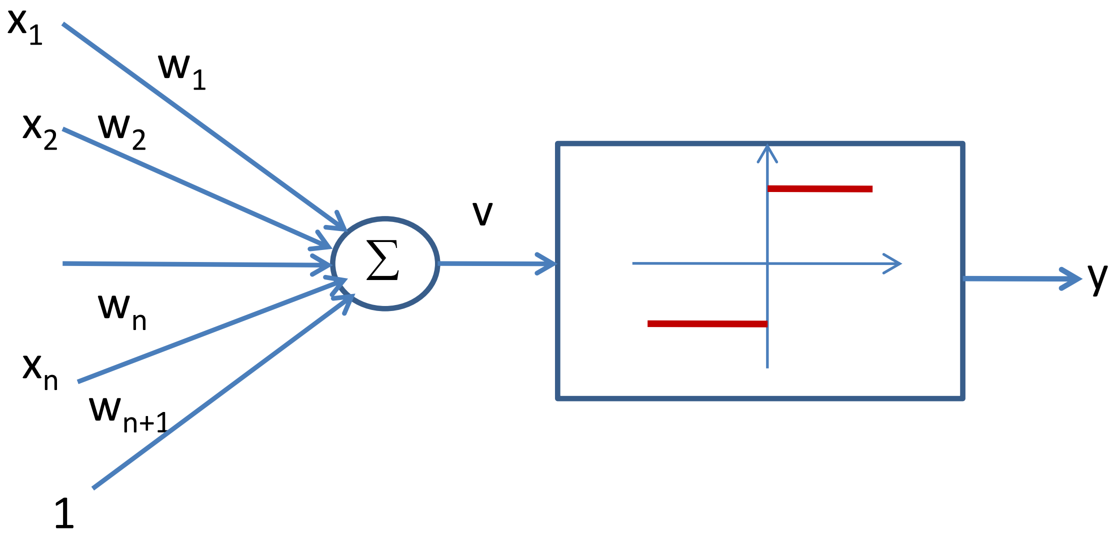
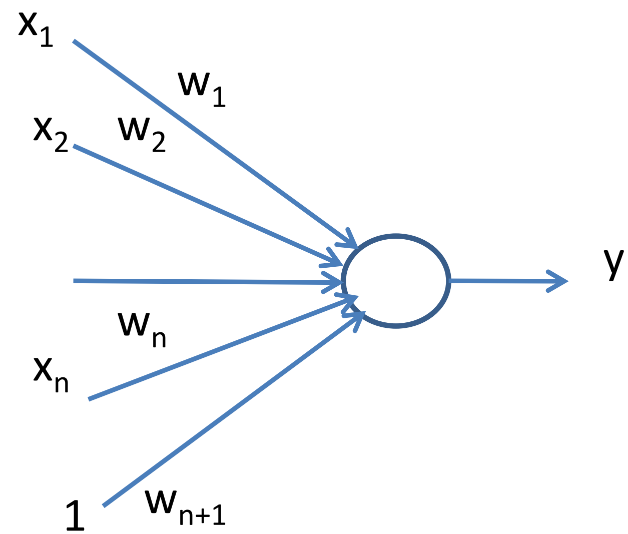
Genlikte Ayrık Algılayıcı-GAA (Perceptron)



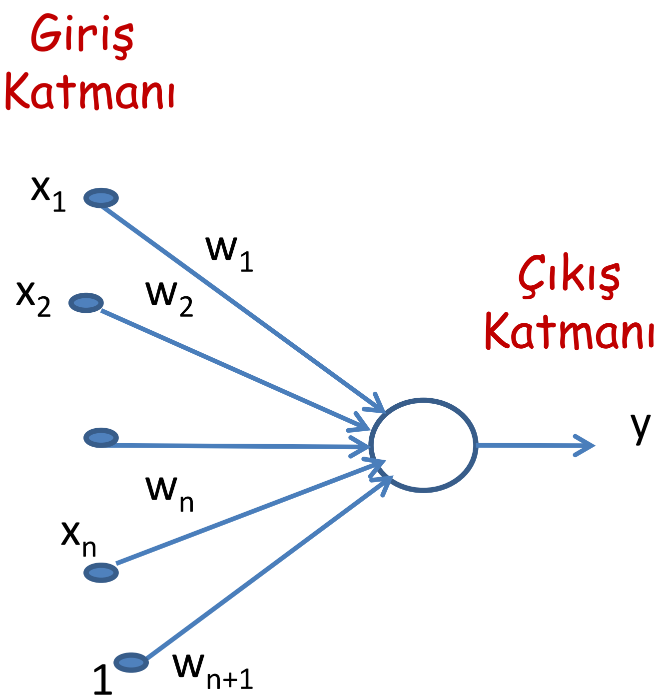


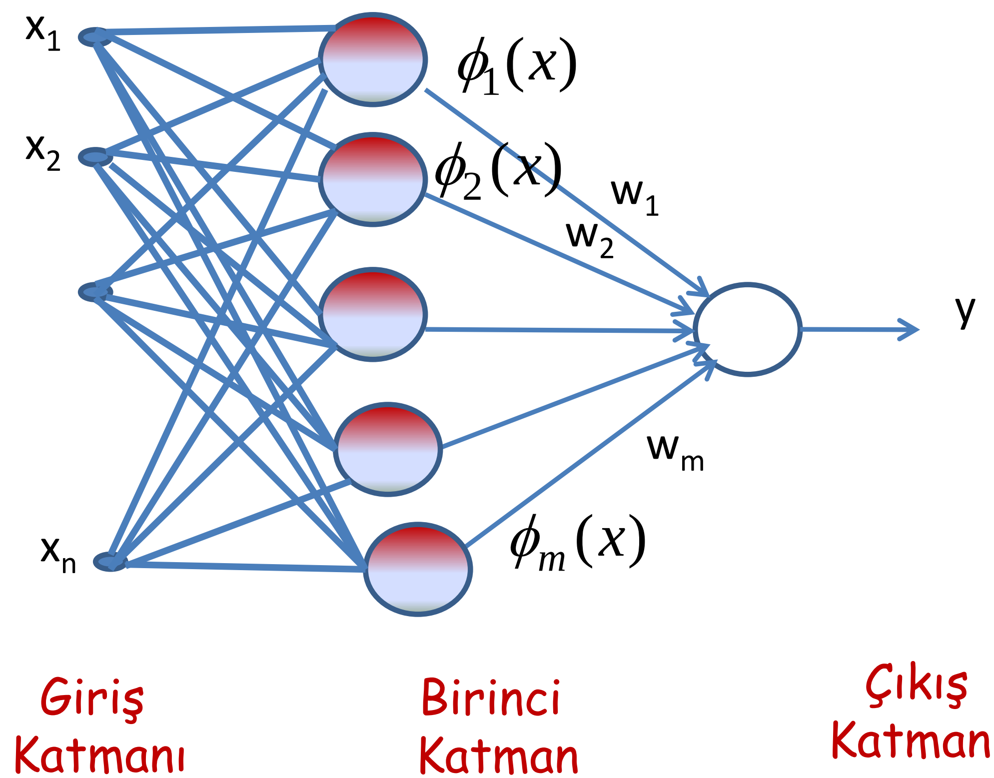
$$v = \begin{bmatrix} w_1 & w_2 & .... & w_n & w_{n+1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \\ 1 \end{bmatrix}$$

$$v = w_1x_1 + w_2x_2 + \ldots + w_nx_n + w_{n+1}1$$

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

Ancak Rosenblatt'ın 1954'de önerdiği yapı bundan farklı





Sabit ağırlıklar, sabit fonksiyonlar

Bağlantı ağırlıkları, eğitim kümesi ile belirlenen tek bir nöron

Genlikte Ayrık Algılayıcı aslında

- Girişlere doğrudan bağlı tek bir nöron değil
- Birinci katman değişmeyen bir yapıya sahip
- Çıkış katmanı, tek bir nörondan oluşan eğitilebilir bir yapı.

Peki Rosenblatt neden birinci katmana gerek duymuş?

Birinci katmanda farklı fonksiyonları oluşturup, öğrenme ile bunlar cinsinden çıkışta istenilen fonksiyonu ifade etmek için

Tanım: <u>Doğrusal ayrıştırılabilir küme</u> (Linearly separable set)

X kümesi R tane X_i alt kümesinden oluşsun. g_i 'ler x'in doğrusal fonksiyonu olmak üzere

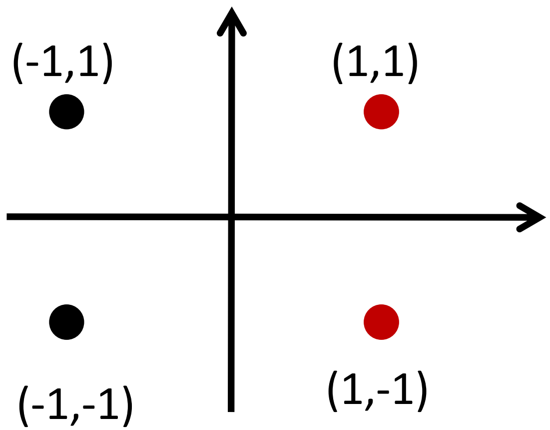
$$g_i(x) > g_j(x) \qquad \forall x \in X_i \qquad i = 1,2,\ldots,R$$

$$j = 1,2,\ldots,R$$

$$i \neq j$$

ise X_i kümeleri doğrusal ayrıştırılabilir kümelerdir.■

Tanımı anlamaya çalışalım...



R=?

$$X_1 = \{(-1,1),(-1,-1)\}$$
$$X_2 = \{(1,1),(1,-1)\}$$

g'leri yazalım

$$g_1 : a_1 x_1 + b_1 x_2 + c_1$$
$$g_2 : a_2 x_1 + b_2 x_2 + c_2$$

Bu iki kümenin doğrusal ayrıştırılabilir olduğunu göstermek için ne yapmalıyız?

$$-a_1 + b_1 + c_1 > -a_2 + b_2 + c_2$$
$$-a_1 - b_1 + c_1 > -a_2 - b_2 + c_2$$

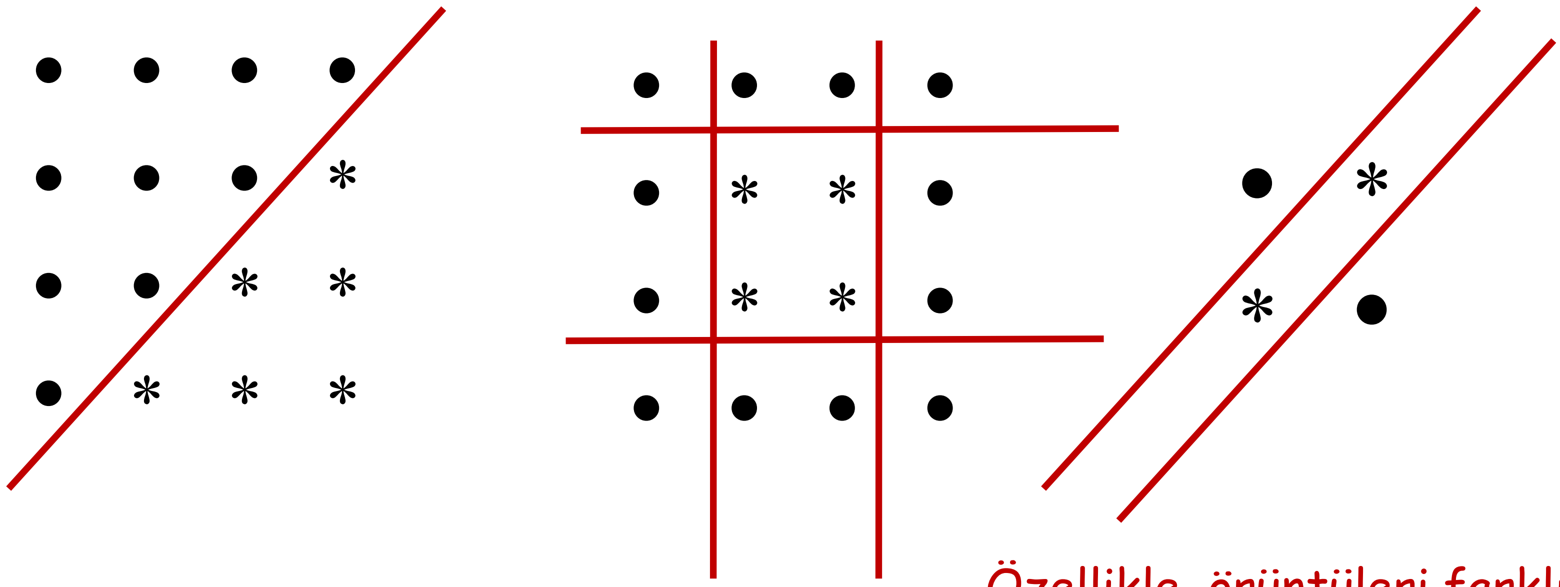
ve

$$a_1 + b_1 + c_1 < a_2 + b_2 + c_2$$
$$a_1 - b_1 + c_1 < a_2 - b_2 + c_2$$

sağlayan a,b ve c'ler bulunmalı

bir çözüm:

$$a_1 = -0.5, b_1 = 0, c_1 = 0,$$
$$a_2 = 0.5, b_2 = 0, c_2 = 0,$$



Özellikle, örüntüleri farklı sınıflara ayıran düzlem olarak kullanılıyor.

Tanım: <u>Karar Düzlemi</u> (Decision Surface) Kümeleri ayıran düzlem. ■

Tek bir nöron ile neler yapılabilir?

Nöron sayısını artırarak ne yapılabilir ne yapılamaz?

Rosenblatt'ın Genlikte Ayrık Algılayıcısında neler oluyor?

Girişler

$$x \in R^n$$

n-boyutlu bir vektör bir örüntüyü temsil ediyor (pattern)

Örüntü uzayı (pattern space)

$$\phi_j(.)$$

$$\phi(.) : R^n \rightarrow R^m$$

Bu dönüşüm genel olarak doğrusal değil

Katman1'in çıkışı

Ne oldu?

$$\hat{y}_j = \phi_j(x) \qquad j = 1,2,\ldots,m$$

$$\hat{y} \in R^m$$

Genlikte Ayrık Algılayıcı için karar düzlemi: $\sum_{i=1}^{m} w_i \hat{y}_i + w_{m+1} = 0$

??????

Genlikte Ayrık Algılayıcı ancak katman 1'in görüntü uzayındaki örüntüleri..........Doğrusal Ayrıştırılabilir.....ise iki sınıfa ayırır.

$$S_1 : \sum_{i=1}^{m} w_i \hat{y}_i + w_{m+1} > 0 \qquad S_2 : \sum_{i=1}^{m} w_i \hat{y}_i + w_{m+1} < 0$$

Rosenblatt'ın Genlikte Ayrık Algılayıcısında neler oluyor?

Girişler

$$x \in R^n$$

n-boyutlu bir vektör bir örüntüyü temsil ediyor (pattern)

Örüntü uzayı (pattern space)

$$\phi_j(.)$$

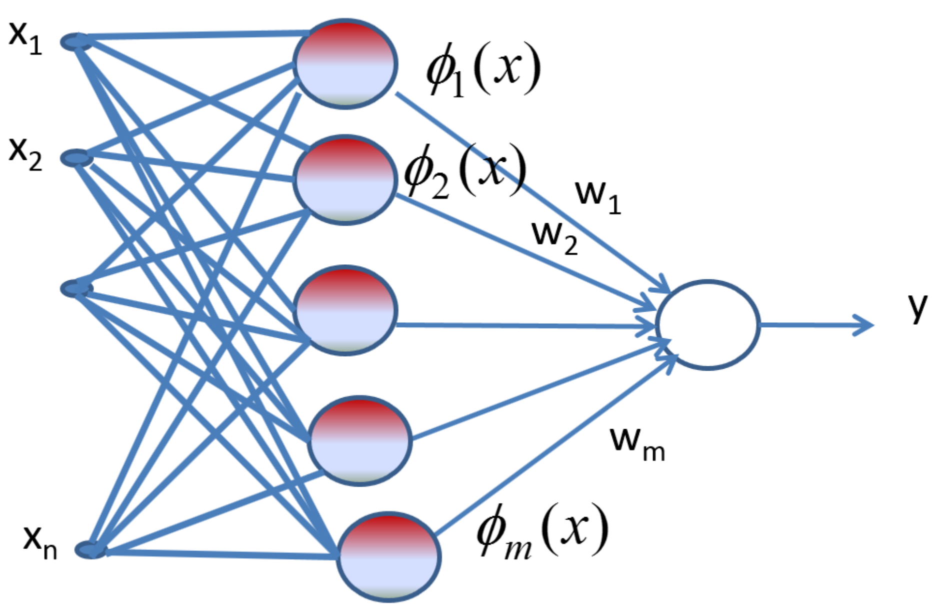
Katman1'in çıkışı

Ne oldu?

$$\hat{y}_j = \phi_j(x) \qquad j = 1,2,\ldots,m$$

$$\hat{y} \in R^m$$

$$\begin{aligned}\hat{y} &= \phi(x)\\ &= \left[\phi_1(x) \quad \phi_2(x) \quad \ldots \quad \phi_m(x)\right]\end{aligned}$$



Soru: Katman 1'de m işlem birimine sahip bir GAA, katman 1 görüntü uzayınındaki P tane örüntüyü 2 sınıfa kaç türlü ayırabilir?

Burada işi ne?

Yanıt:
$$L(P,m)=\begin{cases} 2^P & P \le m+1 \\ 2\sum_{i=0}^{m}\binom{P-1}{i} & P > m+1 \end{cases}$$

Hatırlatma: $\binom{k}{l}=\frac{k!}{l!(k-l)!}$

Soru: Herhangi bir doğrusal karar düzleminin GAA ile hesaplanabilme olasılığı nedir?

Yanıt:
$$\hat{P}_{P,m}=\frac{L(P,m)}{2^P}=\begin{cases} 1 & P \le m+1 \\ 2^{1-P}\sum_{i=0}^{m}\binom{P-1}{i} & P > m+1 \end{cases}$$

$$\hat{P}_{2(m+1),m}=\frac{1}{2}$$

$$\lim_{m\to\infty}\hat{P}_{(2-\varepsilon)(m+1),m}=1$$

m büyük bir sayı ise 2(m+1)'den daha az sayıdaki örüntüyü doğru şekilde sınıflayabilir.

$$\lim_{m\to\infty}\hat{P}_{(2+\varepsilon)(m+1),m}=0$$

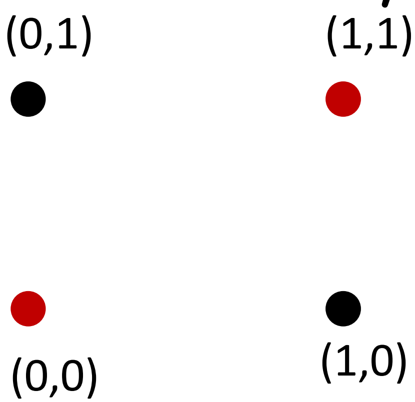
$0 < P < 2(m+1)$

Katman 1 örüntüleri doğrusal ayrıştırılabilecekleri görüntü uzayına taşır. Doğrusal ayrıştırılamayan örüntüleri doğrusal ayrıştırılabilir kılmak iki türlü olasıdır:

(i) m ↗

(ii) P ↘

XOR Genlikte Ayrık Algılayıcı ile ifade etmek:

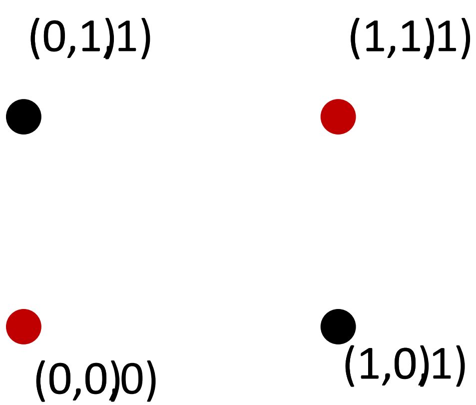


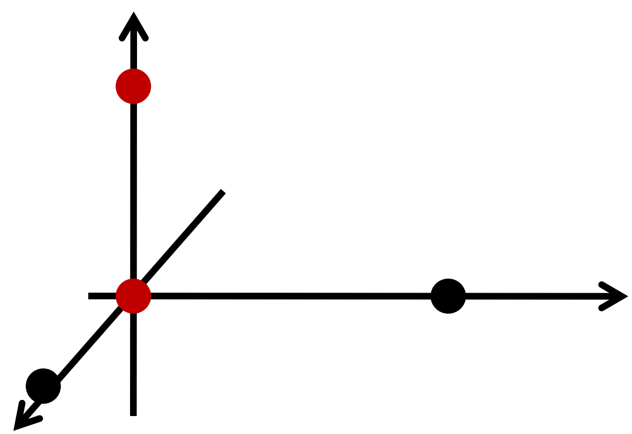
P=4, n=2

İki girişli, tek nöronun kapasitesi: 2*3=6

Kapasite açısından uygun ama doğrusal ayrıştırılabilir değiller

$$\phi(x_1,x_2)=[(-x_1(x_2-1)) \quad (-x_2(x_1-1)) \quad (-(x_1-1)(x_2-1)+1)]$$

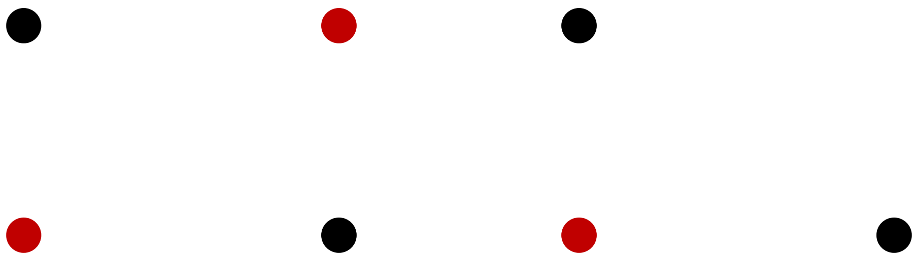




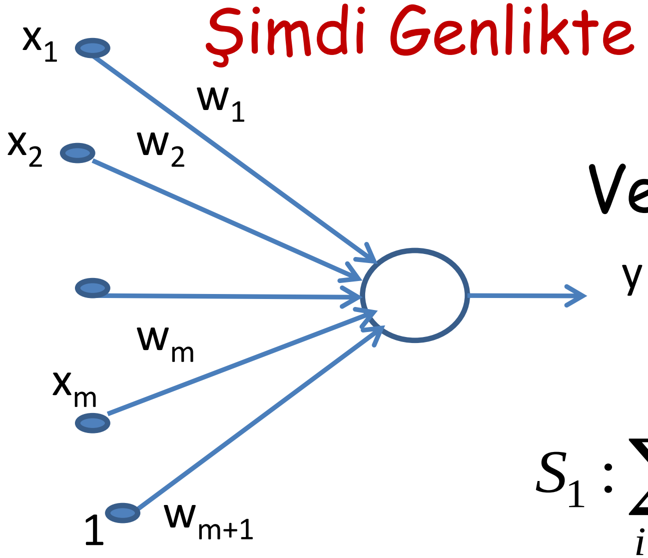
m mi P mi değişti?

$$\phi(x_1, x_2) = [(x_2 - x_1 x_2 - 0.5) \quad (x_1 - x_2 x_1 - 0.5)\,]$$

m mi P mi değişti?



Şimdi Genlikte Ayrık Algılayıcı ile biraz iş yapalım



Verilenler: $\left\{\left(x^k, y_d^k\right)\right\}\Big|_{k=1}^{P}$ Eğitim Kümesi

Amaç: İki sınıfa ayırmak

$$S_1 : \sum_{i=1}^{n} w_i x_i + w_{n+1} > 0 \qquad S_2 : \sum_{i=1}^{n} w_i x_i + w_{n+1} \leq 0$$

Gerçeklenebilme Koşulu: Eğitim kümesi doğrusal ayrıştırılabilir

Eğitim kümesi doğrusal ayrıştırılabilir ise: $\begin{bmatrix} x \\ 1 \end{bmatrix}$

Eğitim kümesi doğrusal ayrıştırılabilir değil ise: $\begin{bmatrix} \phi(x) \\ 1 \end{bmatrix}$

<u>Öğrenme Kuralı:</u>

$$x \in S_1 \qquad w^{\mathrm{T}} x > 0 \qquad w(n+1) = w(n)$$

$$x \in S_2 \qquad w^{\mathrm{T}} x \le 0 \qquad w(n+1) = w(n)$$

$$x \in S_1 \qquad w^{\mathrm{T}} x \le 0 \qquad w(n+1) = w(n) + cx$$

$$x \in S_2 \qquad w^{\mathrm{T}} x > 0 \qquad w(n+1) = w(n) - cx$$

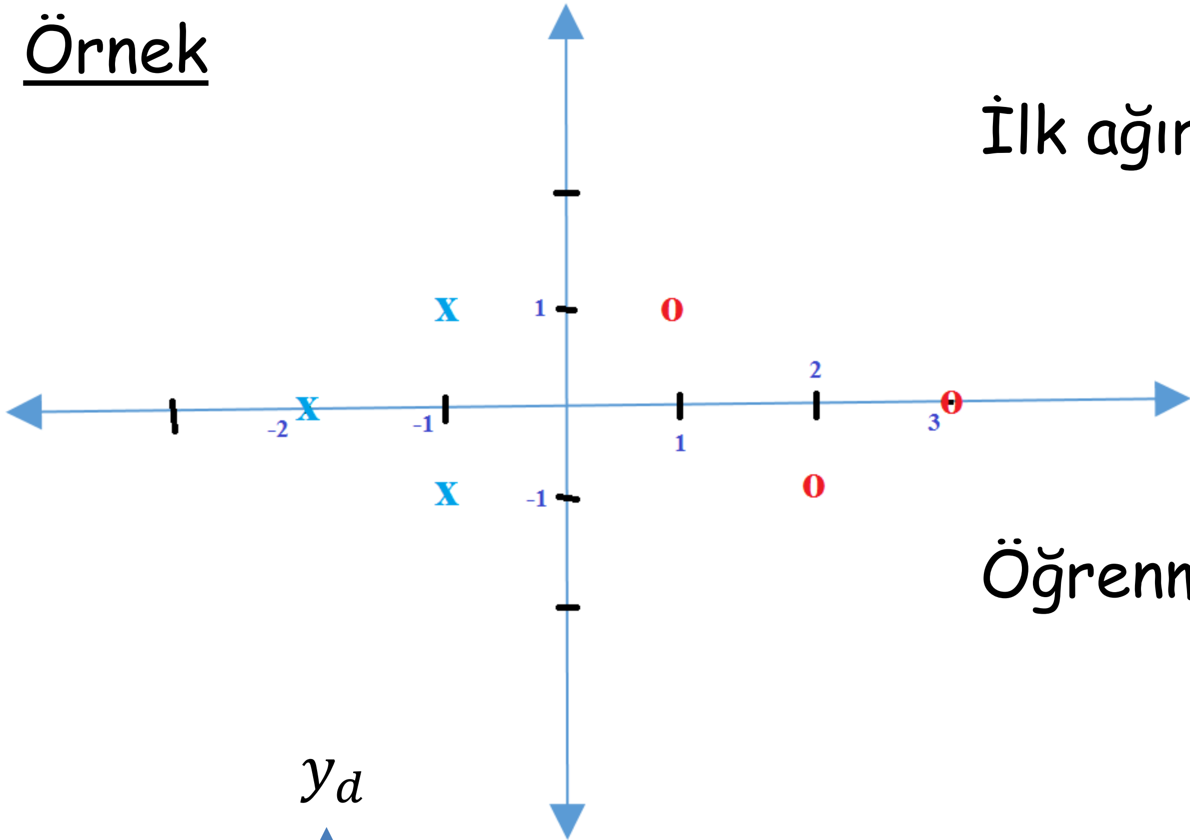
öğrenme hızı <1 olan pozitif bir sayı

$$\rightarrow \quad \Delta w_i = c\frac{1}{2}[y_d - y]x_i$$

$$\Delta w = c\frac{1}{2}[y_d - y]x$$

Örnek

İlk ağırlık $w(0) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$



Öğrenme hızı $c=1$

y_d

$$\left\{ \left(\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, 1 \right), \left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, -1 \right), \left(\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, -1 \right), \left(\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, 1 \right), \left(\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}, -1 \right), \left(\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}, 1 \right) \right\}$$

bais

$$\left\{\left\{\begin{bmatrix}-1\\1\\1\end{bmatrix},1\right\},\left\{\begin{bmatrix}1\\1\\1\end{bmatrix},-1\right\},\left\{\begin{bmatrix}2\\-1\\1\end{bmatrix},-1\right\},\left\{\begin{bmatrix}-1\\-1\\1\end{bmatrix},1\right\},\left\{\begin{bmatrix}3\\0\\1\end{bmatrix},-1\right\},\left\{\begin{bmatrix}-2\\0\\1\end{bmatrix},1\right\}\right\}$$

1. iterasyon

1. örüntü $v_1 = w^T(1)x_1 \Rightarrow \begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}-1\\1\\1\end{bmatrix} = 1 \xrightarrow{\varphi} y_1 = 1$

$$w(2) = w(1) + \frac{1}{2}c(y_{1d} - y_1)x_1 \;\Rightarrow\; w(2) = \begin{bmatrix}1\\1\\1\end{bmatrix} + \frac{1}{2}\,1(1-1)\begin{bmatrix}-1\\1\\1\end{bmatrix} = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

2. örüntü $v_2 = w^T(2)x_2 \Rightarrow \begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix} = 3 \xrightarrow{\varphi} y_2 = 1$

$$w(3) = w(2) + \frac{1}{2}c(y_{2d} - y_2)x_2 \;\Rightarrow\; w(3) = \begin{bmatrix}1\\1\\1\end{bmatrix} + \frac{1}{2}\,1(-1-1)\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

$$\left\{\left(\begin{bmatrix}-1\\1\\1\end{bmatrix},1\right),\left(\begin{bmatrix}1\\1\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}2\\-1\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}-1\\-1\\1\end{bmatrix},1\right),\left(\begin{bmatrix}3\\0\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}-2\\0\\1\end{bmatrix},1\right)\right\}$$

3. örüntü $\quad v_3 = w^T(3)x_3 \Rightarrow [0 \quad 0 \quad 0]\begin{bmatrix}2\\-1\\1\end{bmatrix} = 0 \xrightarrow{\varphi} y_3 = -1$

$$w(4) = w(3) + \frac{1}{2}c(y_{3d} - y_3)x_3 \;\Rightarrow\; w(4) = \begin{bmatrix}0\\0\\0\end{bmatrix} + \frac{1}{2}\,1(-1+1)\begin{bmatrix}2\\-1\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

4. örüntü $\quad v_4 = w^T(4)x_4 \Rightarrow [0 \quad 0 \quad 0]\begin{bmatrix}-1\\-1\\1\end{bmatrix} = 0 \xrightarrow{\varphi} y_4 = -1$

$$w(5) = w(4) + \frac{1}{2}c(y_{4d} - y_4)x_4 \;\Rightarrow\; w(5) = \begin{bmatrix}0\\0\\0\end{bmatrix} + \frac{1}{2}\,1(1+1)\begin{bmatrix}-1\\-1\\1\end{bmatrix} = \begin{bmatrix}-1\\-1\\1\end{bmatrix}$$

$$\left\{\left(\begin{bmatrix}-1\\1\\1\end{bmatrix},1\right),\left(\begin{bmatrix}1\\1\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}2\\-1\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}-1\\-1\\1\end{bmatrix},1\right),\left(\begin{bmatrix}3\\0\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}-2\\0\\1\end{bmatrix},1\right)\right\}$$

5. örüntü $\quad v_5 = w^T(5)x_5 \Rightarrow [-1 \quad -1 \quad 1]\begin{bmatrix}3\\0\\1\end{bmatrix} = -2 \xrightarrow{\varphi} y_5 = -1$

$$w(6) = w(5) + \frac{1}{2}c(y_{5d} - y_5)x_5 \;\Rightarrow\; w(6) = \begin{bmatrix}-1\\-1\\1\end{bmatrix} + \frac{1}{2}\,1(-1+1)\begin{bmatrix}2\\-1\\1\end{bmatrix} = \begin{bmatrix}-1\\-1\\1\end{bmatrix}$$

6. örüntü $\quad v_6 = w^T(6)x_6 \Rightarrow [-1 \quad -1 \quad 1]\begin{bmatrix}-2\\0\\1\end{bmatrix} = 3 \xrightarrow{\varphi} y_6 = 1$

$$w(7) = w(6) + \frac{1}{2}c(y_{6d} - y_6)x_6 \;\Rightarrow\; w(7) = \begin{bmatrix}-1\\-1\\1\end{bmatrix} + \frac{1}{2}\,1(1-1)\begin{bmatrix}-2\\0\\1\end{bmatrix} = \begin{bmatrix}-1\\-1\\1\end{bmatrix}$$

1. İterasyon tamamlandı. Doğru sınıflandırılan örüntü sayısı: 4

Toplam örüntü sayısı: 6

$$\left\{\left(\begin{bmatrix}-1\\1\\1\end{bmatrix},1\right),\left(\begin{bmatrix}1\\1\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}2\\-1\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}-1\\-1\\1\end{bmatrix},1\right),\left(\begin{bmatrix}3\\0\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}-2\\0\\1\end{bmatrix},1\right)\right\}$$

2. iterasyon

1. örüntü $v_1 = w^T(7)x_1 \Rightarrow \begin{bmatrix}-1 & -1 & 1\end{bmatrix}\begin{bmatrix}-1\\1\\1\end{bmatrix} = 1 \xrightarrow{\varphi} y_1 = 1$

$$w(8) = w(7) + \frac{1}{2}c(y_{1d} - y_1)x_1 \;\Rightarrow\; w(8) = \begin{bmatrix}-1\\-1\\1\end{bmatrix} + \frac{1}{2}\,1(1-1)\begin{bmatrix}-1\\1\\1\end{bmatrix} = \begin{bmatrix}-1\\-1\\1\end{bmatrix}$$

2. örüntü $v_2 = w^T(8)x_2 \Rightarrow \begin{bmatrix}-1 & -1 & 1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix} = -1 \xrightarrow{\varphi} y_2 = -1$

$$w(9) = w(8) + \frac{1}{2}c(y_{2d} - y_2)x_2 \;\Rightarrow\; w(9) = \begin{bmatrix}-1\\-1\\1\end{bmatrix} + \frac{1}{2}\,1(-1+1)\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}-1\\-1\\1\end{bmatrix}$$

$$\left\{\left(\begin{bmatrix}-1\\1\\1\end{bmatrix},1\right),\left(\begin{bmatrix}1\\1\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}2\\-1\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}-1\\-1\\1\end{bmatrix},1\right),\left(\begin{bmatrix}3\\0\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}-2\\0\\1\end{bmatrix},1\right)\right\}$$

3. örüntü $v_3 = w^T(9)x_3 \Rightarrow \begin{bmatrix}-1 & -1 & 1\end{bmatrix}\begin{bmatrix}2\\-1\\1\end{bmatrix} = 0 \xrightarrow{\varphi} y_3 = -1$

$$w(10) = w(9) + \frac{1}{2}c(y_{3d} - y_3)x_3 \Rightarrow w(10) = \begin{bmatrix}-1\\-1\\1\end{bmatrix} + \frac{1}{2}\,1(-1+1)\begin{bmatrix}2\\-1\\1\end{bmatrix} = \begin{bmatrix}-1\\-1\\1\end{bmatrix}$$

4. örüntü $v_4 = w^T(10)x_2 \Rightarrow \begin{bmatrix}-1 & -1 & 1\end{bmatrix}\begin{bmatrix}-1\\-1\\1\end{bmatrix} = 3 \xrightarrow{\varphi} y_4 = 1$

$$w(11) = w(10) + \frac{1}{2}c(y_{4d} - y_4)x_4 \Rightarrow w(11) = \begin{bmatrix}-1\\-1\\1\end{bmatrix} + \frac{1}{2}\,1(1-1)\begin{bmatrix}-1\\-1\\1\end{bmatrix} = \begin{bmatrix}-1\\-1\\1\end{bmatrix}$$

$$\left\{\left(\begin{bmatrix}-1\\1\\1\end{bmatrix},1\right),\left(\begin{bmatrix}1\\1\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}2\\-1\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}-1\\-1\\1\end{bmatrix},1\right),\left(\begin{bmatrix}3\\0\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}-2\\0\\1\end{bmatrix},1\right)\right\}$$

5. örüntü $\quad v_5 = w^T(11)x_5 \Rightarrow [-1 \quad -1 \quad 1]\begin{bmatrix}3\\0\\1\end{bmatrix} = -2 \xrightarrow{\varphi} y_5 = -1$

$$w(12) = w(11) + \frac{1}{2}c(y_{5d} - y_5)x_5 \;\Rightarrow\; w(12) = \begin{bmatrix}-1\\-1\\1\end{bmatrix} + \frac{1}{2}\,1(-1+1)\begin{bmatrix}3\\0\\1\end{bmatrix} = \begin{bmatrix}-1\\-1\\1\end{bmatrix}$$

6. örüntü $\quad v_6 = w^T(12)x_6 \Rightarrow [-1 \quad -1 \quad 1]\begin{bmatrix}-2\\0\\1\end{bmatrix} = 3 \xrightarrow{\varphi} y_6 = 1$

$$w(13) = w(12) + \frac{1}{2}c(y_{6d} - y_6)x_6 \;\Rightarrow\; w(13) = \begin{bmatrix}-1\\-1\\1\end{bmatrix} + \frac{1}{2}\,1(1-1)\begin{bmatrix}-2\\0\\1\end{bmatrix} = \begin{bmatrix}-1\\-1\\1\end{bmatrix}$$

2. İterasyon tamamlandı. Doğru sınıflandırılan örüntü sayısı: 6

Toplam örüntü sayısı: 6

Tüm örüntüler doğru sınıflandırıldığı için ağın eğitim aşaması tamamlanmıştır.

Ağın Test Aşaması

$$\left\{\left\{\begin{bmatrix}2\\2\\1\end{bmatrix}, -1\right\}, \left\{\begin{bmatrix}-2\\-1\\1\end{bmatrix}, 1\right\}\right\}$$

1. örüntü $v_1 = w^T(13)x_1 \Rightarrow [-1 \quad -1 \quad 1]\begin{bmatrix}2\\2\\1\end{bmatrix} = -3 \xrightarrow{\varphi} y_1 = -1$

$y_{d1} = -1$ ✓

2. örüntü $v_2 = w^T(13)x_2 \Rightarrow [-1 \quad -1 \quad 1]\begin{bmatrix}-2\\-1\\1\end{bmatrix} = 4 \xrightarrow{\varphi} y_1 = 1$

$y_{d2} = 1$ ✓