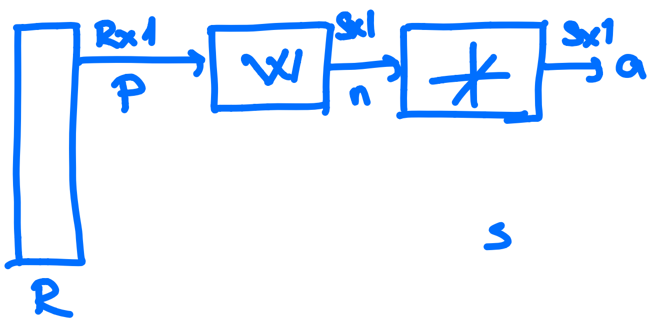
DANIŞMANLI HEBBIAN ÖĞRENMESİ

Hebb kuralı ilk YSA öğrenme kuralıdır. Donald Hebb tarafından 1949 yılında geliştirilmiştir. İnsan beynindeki sinaptik değişimleri temsil edebilecek bir mekanizma olarak önerilmiştir.

Hebb Postulatı:
"A hücresinin bir aksonu, bir B hücresini uyaracak kadar yakın olduğunda ve tekrar tekrar veya ısrarla ateşlemede yer aldığında, B hücresini ateşleyen hücrelerden biri olarak A'nın etkinliğini sağlayacak şekilde bir veya iki hücrede bir miktar büyüme süreci veya metabolik değişim gerçekleşir. , artırılır. "

Doğrusal İlişkilendirici:



$$a = purelin(Wp) = Wp$$

$$a_i = \sum_{j=1}^{R} w_{ij} p_j$$

İlişkisel Bellek:

$$\{p_1, t_1\}, \{p_2, t_2\}, \ldots\ldots, \{p_Q, t_Q\}$$

$$p = p_q \rightarrow a = t_q \qquad q = 1, 2, \ldots\ldots Q$$

$$p = p_q + \delta \rightarrow a = t_q + \varepsilon$$

Hebb Kuralı:

$$w_{ij}^{yeni} = w_{ij}^{eski} + \alpha f_i(a_{iq})\, g_j(p_{jq})$$

p_{jq} : q giriş vektörü p_q'nun j. elemanıdır.

a_{iq} : q. giriş vektöründen, i. elemanının çıkışıdır.

$$\underbrace{w_{ij}^{yeni} = w_{ij}^{eski} + \alpha\, a_{iq}\, p_{jq}}_{\text{danışmansız Hebb kuralı}}$$

Danışmanlı Hebb Kuralı:

$$w_{ij}^{yeni} = w_{ij}^{eski} + t_{iq}\, p_{jq}$$

t_{iq} : t_q hedef vektörünün i. elemanıdır.

Vektör notasyonu:

$$W^{yeni} = W^{eski} + t_q p_q^T$$

$$W = t_1 p_1^T + t_2 p_2^T + \dots + t_q p_q^T$$

$$= \sum_{q=1}^{Q} t_q p_q^T$$

$$W = [t_1 \, t_2 \dots t_Q] \begin{bmatrix} p_1^T \\ p_2^T \\ \vdots \\ p_Q^T \end{bmatrix} = TP^T$$

$$T = [t_1 \; t_2 \; \dots \; t_Q] \; , \; P = [p_1 p_2 \dots p_Q]$$

Performans Analizi:

$p_q \rightarrow$ ortonormal (ortogonal ve birim uzunluklu)

Eğer p_k giriş ise, ağın çıkışı:

$$a = W.p_k = \left(\sum_{q=1}^{Q} t_q p_q^T \right) p_k$$

$$= \sum_{q=1}^{Q} t_q (p_q^T p_k)$$

p_q ortonormal olduğundan dolayı

$$(p_q^T p_k) = 1 \qquad q = k$$

$$= 0 \qquad q \neq k$$

$$a = W p_k = t_k$$

Yorum: Ağın çıkışı hedef çıkışa eşittir.
Bu durum giriş prototip vektörü ortonormal ise Hebb kuralı her giriş için doğru çıkışı üretecektir.

$$a = W p_k = t_k + \underbrace{\boxed{\sum_{q \neq k} t_q \left(p_q^T p_k\right)}}_{\text{Hata}}$$

ÖRNEK:

$$\left.\begin{array}{ll} p_1 = \begin{bmatrix} 0.5 \\ -0.5 \\ 0.5 \\ -0.5 \end{bmatrix} & t_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ p_2 = \begin{bmatrix} 0.5 \\ 0.5 \\ -0.5 \\ -0.5 \end{bmatrix} & t_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{array}\right\} \text{Her iki vektörün ortonormal}$$

Ağırlık vektörü

$$W = T.P^T = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 0.5 & -0.5 & 0.5 & -0.5 \\ 0.5 & 0.5 & -0.5 & -0.5 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix}$$

$$Wp_1 = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0.5 \\ -0.5 \\ 0.5 \\ -0.5 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$Wp_2 = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \\ -0.5 \\ -0.5 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

ÖRNEK:

$$p_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \qquad p_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \Rightarrow$$ ortogonal değiller

Eğer girişleri normalize eder ve arzu edilen çıkışlarımızı -1 ve 1 olarak belirler isek:

$$p_1 = \begin{bmatrix} 0.5774 \\ -0.5774 \\ -0.5774 \end{bmatrix} \qquad t_1 = \begin{bmatrix} -1 \end{bmatrix}$$

$$p_2 = \begin{bmatrix} 0.5774 \\ 0.5774 \\ -0.5774 \end{bmatrix} \quad t_2 = [1]$$

Ağırlık matrisi:

$$W = TP^T = [-1 \quad 1] \begin{bmatrix} 0.5774 & -0.5774 & -0.5774 \\ 0.5774 & 0.5774 & -0.5774 \end{bmatrix}$$

$$= [0 \quad 1.1548 \quad 0]$$

$$Wp_1 = [0 \quad 1.1548 \quad 0] \begin{bmatrix} 0.5774 \\ -0.5774 \\ -0.5774 \end{bmatrix} = [-0.6668]$$

$$Wp_2 = [0 \quad 1.1548 \quad 0] \begin{bmatrix} 0.5774 \\ 0.5774 \\ -0.5774 \end{bmatrix} = [0.6668]$$

Hebb Öğrenme Kuralının Varyasyonları:

$$W^{yeni} = W^{eski} + t_q p_q^T$$

$\alpha \rightarrow$ öğrenme oranı

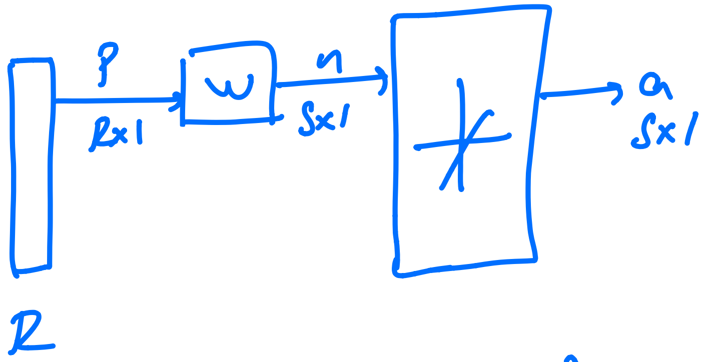
$$W^{yeni} = W^{eski} + \alpha t_q p_q^T$$

$$W^{yeni} = W^{eski} + \alpha t_q p_q^T - \gamma W^{eski}$$

$$= (1-\gamma) W^{eski} + \alpha t_q q_q^T$$

$$\gamma < 1$$

Lineer İlişkilendirici:



$$a = purelin(Wp)$$

Hebb Kuralı: $w_{ij}^{yeni} = w_{ij}^{eski} + t_{qi} p_{qj}$

$$W = t_1 p_1^T + t_2 p_2^T + \dots + t_q p_q^T$$

$$W = \begin{bmatrix} t_1 & t_2 & \dots & t_q \end{bmatrix} \begin{bmatrix} p_1^T \\ p_2^T \\ \vdots \\ p_q^T \end{bmatrix} = TP^T$$

Hebb Kuralının Varyasyonları

Filtrelenen öğrenme

$$W^{yeni} = (1-\gamma) W^{eski} + \alpha t_q p_q^T$$

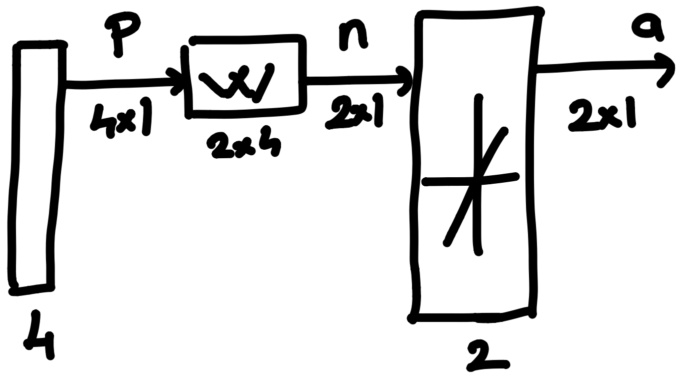
Delta Kuralı

$$W^{yeni} = W^{eski} + \alpha (t_q - a_q) p_q^T$$

Danışmanlı
Hebb

$$w^{yeni} = w^{eski} + \alpha a_q p_q^T$$

ÖRNEK:



$$p_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \quad t_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$p_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \quad t_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$P = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 1 & -1 \\ -1 & -1 \end{bmatrix} \qquad T = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$W = \begin{bmatrix} t_1 & t_2 & \dots & t_q \end{bmatrix} \begin{bmatrix} p_1^T \\ p_2^T \\ \\ p_q^T \end{bmatrix} = TP^T$$

$$W^h = TP^T = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & 0 & -2 \\ 0 & 2 & -2 & 0 \end{bmatrix}$$