

- Bir Yapay Sinir Ağı Tanımı** (Alexander, Morton 1990)

Yapay sinir ağı, basit işlemci ünitelerinden oluşmuş, çok yoğun, paralel ve dağılmış düzende çalışan bir işlemcidir.

Deneysel bilgiyi depolama ve kullanıma sunma özelliğine sahiptir.

Beyni iki şekilde andırır:

- 1) Ağ, bilgiyi ortamdan öğrenme yolu ile elde eder.
- 2) Gerekli bilgiyi depolama için basit işlemci ünitelerin arasındaki bağlantıları kullanır.

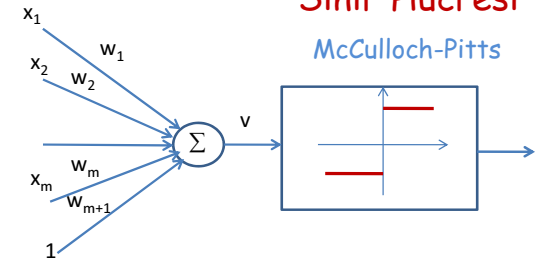
Sinaptik
ağırlıklar

Sinir
hücresi

1

Sinir Hücresi

McCulloch-Pitts



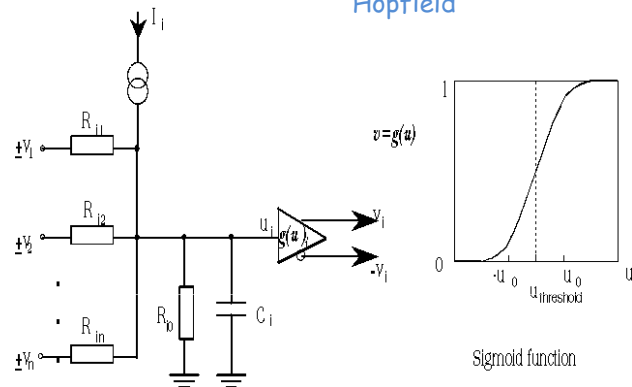
$$v = [w_1 \ w_2 \ \dots \ w_m \ w_{m+1}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

$$v = w_1 x_1 + w_2 x_2 + \dots + w_m x_m + w_{m+1} 1$$

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

2

Hopfield



Sigmoid function

$$\frac{du_i}{dt} = -\frac{u_i}{\tau_i} + \sum_{j=1}^n T_{ij} v_j + I_i$$

3

Aktivasyon Fonksiyonu

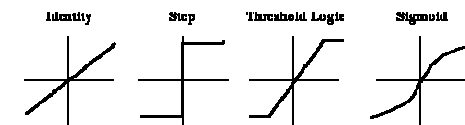
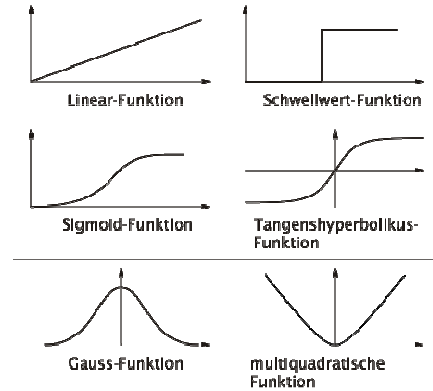


Figure 2: Various activation functions used in connectionist networks

<http://intsys.mgt.qub.ac.uk/notes/image/activefn.gif>

4

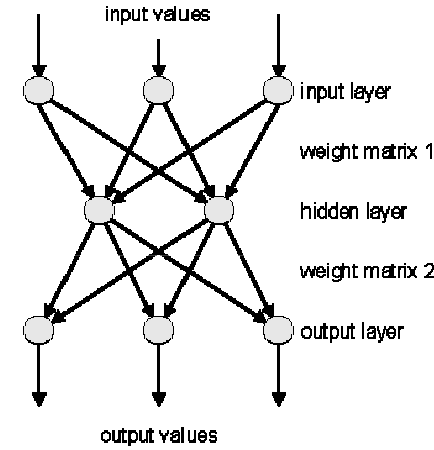
Aktivasyon Fonksiyonu



http://www.irt.rwth-aachen.de/uploads/pics/KNNFunktionen_01.png

5

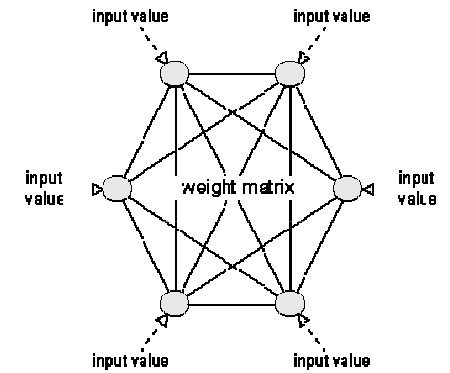
Ağ Yapıları



İleri yol ağı-feedforward

Çok katmanlı Ağ-Multilayer Perceptron

<http://fbim.fh-regensburg.de/~saj39122/jfroehl/diplom/e-12-text.html>

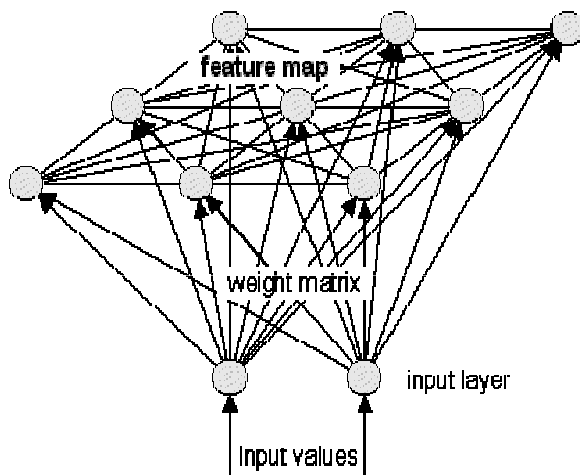


Tam bağlaşımlı ağ-reccurent

Hopfield Ağı

6

Ağ Yapıları



Hücresel Ağ
Kohonen Ağı

<http://fbim.fh-regensburg.de/~saj39122/jfroehl/diplom/e-12-text.html>

7

"Bilgi"nin Gösterimi"Bilgi"

İnsan
veya
Makina

Yorumlama
Öngörme
Uygun yanıt
verme

Depolanmış
enformasyon
veya
model

Nasıl anlayacağız? Nasıl gösterimi oluşturacağız?

- Kurallar:
- (1) Benzer sınıflardan benzer girişler ağda benzer gösterimler oluşturmali ve böylece aynı kategoriye ait olarak sınıflanmalı,
 - (2) Farklı sınıflara ayrılacak nesnelere, ağda çok farklı gösterimler atanmalı,
 - (3) Belirli bir özellik önemli ise ağda onun gösterimi ile görevlendirilen hücre sayısı daha fazla olmalı,

8

Benzerliğin bir ölçütü - Norm

V vektör uzayı olmak üzere, aşağıdaki dört özelliği sağlayan fonksiyon $\|\cdot\| : V \rightarrow \mathbb{R}$ normdur

$$\|x\| \geq 0$$

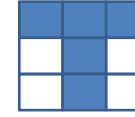
$$\|x\| = 0 \Leftrightarrow x = 0$$

$$\|\alpha x\| = |\alpha| \|x\|$$

$$\|x + y\| \leq \|x\| + \|y\|$$

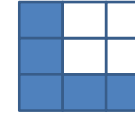
9

Gösterim için bir yol



İşlem uygulayacağımıza göre nasıl ifade edebiliriz?

Bir T harfi



Bir L harfi

$$L = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$T =$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

10

Öğrenme İşlemi

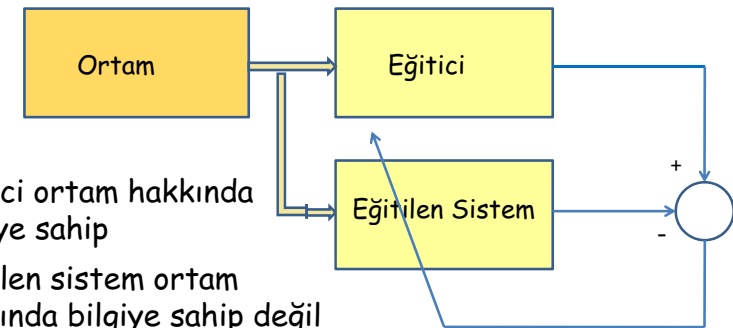
Ağın içinde bulunduğu ortamdan etkilenecek parametrelerini değiştirmesi işlemi öğrenmedir.

Nasıl etkilenebilir?

Öğrenme şekli, parametrelerin nasıl değiştirildiği ile belirlenir.

11

Eğitici Öğrenme



Eğitici ortam hakkında bilgiye sahip

Eğitilen sistem ortam hakkında bilgiye sahip değil

Eğitilen sisteme ilişkin

eğitim kümesinde içerilen bilgi ve aracılığı ile değiştiriliyor

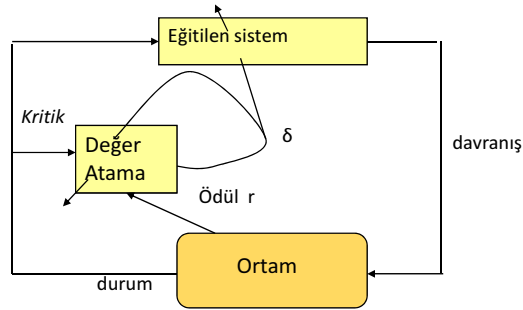
Eğiticinin ortam hakkında sahip olduğu bilgi, eğitilen sisteme aktarılıyor

Burada rahatsız edici bir şey var, ne?

12

Eğitici-siz Öğrenme

• Pekiştirmeli Öğrenme (reinforcement learning)

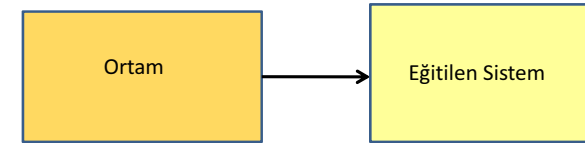


Öğrenme işleminin her adımında istenilen yanıtı sağlayan bir eğitici yok

Eğitilen sistem, sonuçta elde edilecek yanıtı erişmek için gerekli davranışı eleştiriyi gözönünde tutarak bulmak bulmak zorunda

13

• Özdüzenlemeli öğrenme (self-organizing)

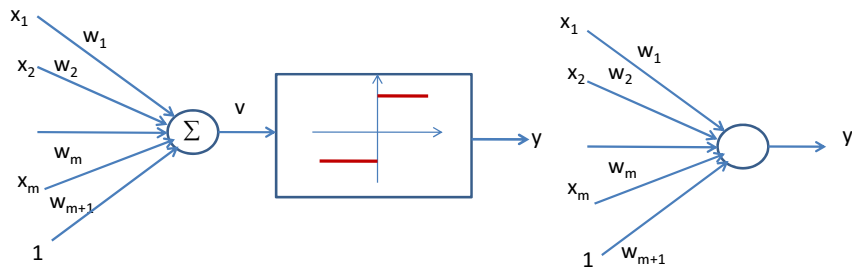


Bilgiye sahip ya da eleştiride bulunan bir eğitici yok

Eğitilen sistem girişlerin istatistiksel dağılımını belirledikten sonra sınıflamayı oluşturuyor.

14

Genlikte Ayırık Algılayıcı-GAA (Perceptron)



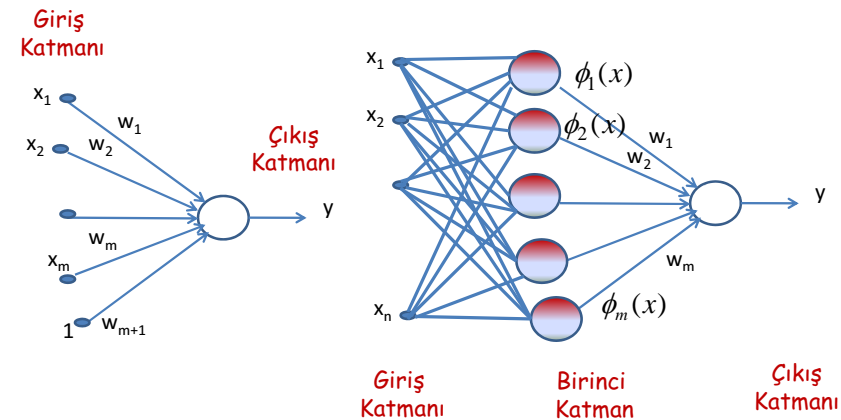
$$v = [w_1 \ w_2 \ \dots \ w_m \ w_{m+1}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \\ 1 \end{bmatrix}$$

$$v = w_1 x_1 + w_2 x_2 + \dots + w_m x_m + w_{m+1} 1$$

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

15

Ancak Rosenblatt'ın 1954'de önerdiği yapı bundan farklı



Sabit ağırlıklar, sabit fonksiyonlar

Bağlantı ağırlıkları, eğitim kümesi ile belirlenen tek bir nöron

16

Genlikte Ayırık 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

Tanım: Doğrusal ayrıştırılabilir küme (Linearly separable set)

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

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

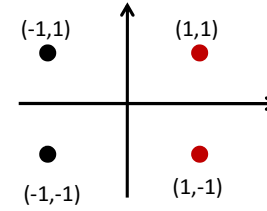
$$j = 1, 2, \dots, R$$

$$i \neq j$$

ise X_i kümeleri doğrusal ayrıştırılabilir kümelerdir. ■

17

Tanımı anlamaya çalışalım...



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

R=?

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

g'leri yazalım

$$g_1 : a_1x_1 + b_1x_2 + c_1$$

$$g_2 : a_2x_1 + b_2x_2 + c_2$$

Bu iki kümenin doğrusal ayrıştırılabilir olduğunu göstermek için ne yapmalıyız?

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

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

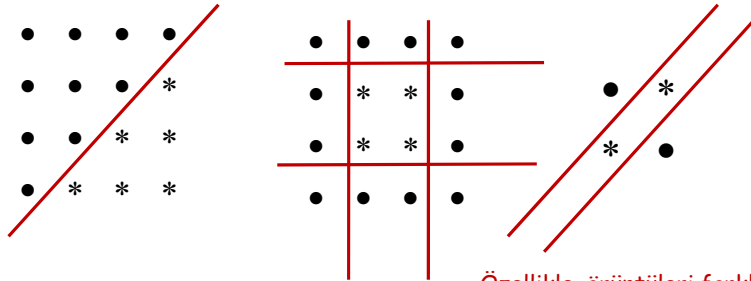
$$-a_1 - b_1 + c_1 > -a_2 - b_2 + c_2$$

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

sağlayan a,b ve c'ler bulunmalı

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

18



Özellikle, örüntüleri farklı sınıflara ayıran düzlem olarak kullanılıyor.

Tanım: Karar Düzlemi (Decision Surface) Kümeleri ayıran düzlem. ■

Tek bir nöron ile neler yapılabilir?

Nöron sayısını artırarak ne yapılabilir ne yapılamaz?

19

Rosenblatt'ın Genlikte Ayırık Algılayıcısında neler oluyor?

Girişler

$$x \in R^n$$

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

Örüntü uzayı (pattern space)

$$\phi_j(\cdot)$$

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

Katman1'in çıkışı

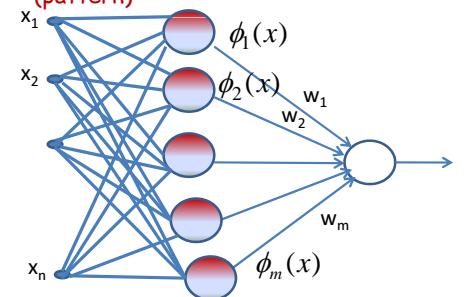
Ne oldu?

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

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

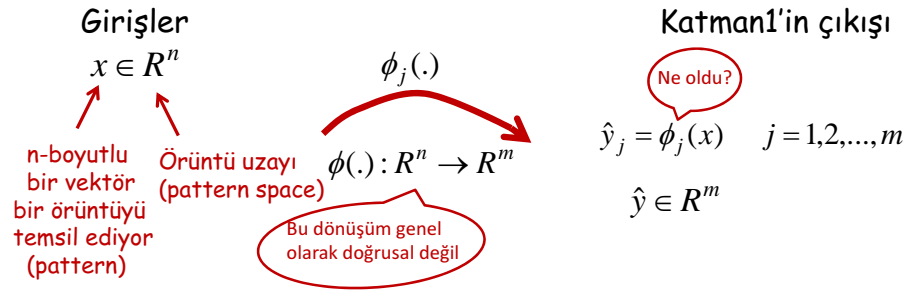
$$\hat{y} = \phi(x)$$

$$= [\phi_1(x) \quad \phi_2(x) \quad \dots \quad \phi_m(x)]$$



20

Rosenblatt'ın Genlikte Ayırık Algılayıcısında neler oluyor?



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

Genlikte Ayırık Algılayıcı ancak katman 1'in görüntü uzayındaki örüntüleri.....ise iki sınıfa ayırır.

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

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

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

Burada işi ne? Hatırlatma: $\binom{k}{l} = \frac{k!}{l!(k-l)!}$

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

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

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

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

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

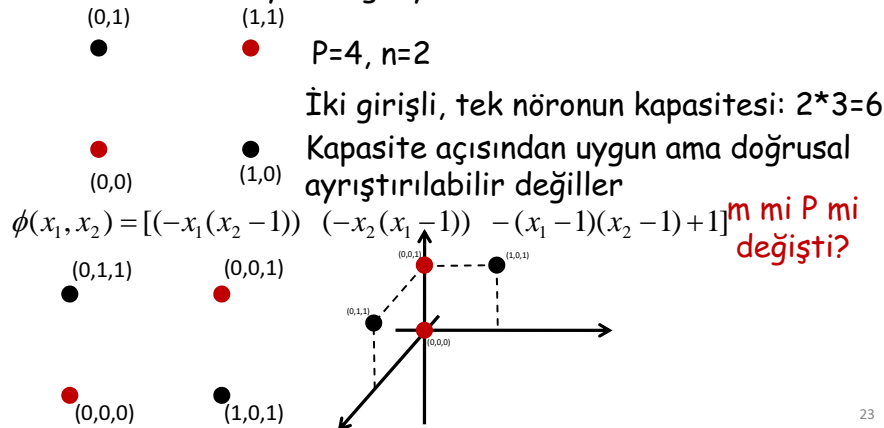
m büyük bir sayı ise $2(m+1)$ 'den daha az sayıda örüntüyü doğru şekilde sınıflayabilir.

$$0 < \hat{P} < 2(m+1)$$

Katman 1 örüntüleri doğrusal ayrıştırılabilecekleri görüntü uzayına taşır. Doğrusal ayrıştırılamayan örüntüleri doğrusal ayrıştırılabilir kılmak iki türlü olasıdır: (i) m ↑

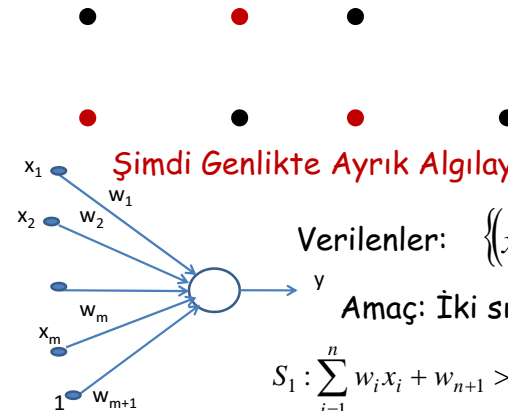
(ii) P ↓

XOR Genlikte Ayırık Algılayıcı ile ifade etmek:



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

m mi P mi değişti?



Şimdi Genlikte Ayırık Algılayıcı ile biraz iş yapalım

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

Amaç: İki sınıfa ayırmak

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

Gerçeklemlenebilir Koşulu: Eğitim kümesi doğrusal ayrıştırılabilir

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

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

Öğrenme Kuralı:

$$x \in S_1 \quad w^T x > 0 \quad w(k+1) = w(k)$$

$$x \in S_2 \quad w^T x \leq 0 \quad w(k+1) = w(k)$$

$$x \in S_1 \quad w^T x \leq 0 \quad w(k+1) = w(k) + cx$$

$$x \in S_2 \quad w^T x > 0 \quad w(k+1) = w(k) - cx$$

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

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

öğrenme hızı <1
olan pozitif bir sayı