

Derin Öğrenme Deep Learning

Hazırlayan: M. Ali Akcayol
Gazi Üniversitesi
Bilgisayar Mühendisliği Bölümü

Bu dersin sunumları, "Ian Goodfellow, Yoshua Bengio and Aaron Courville, Deep Learning, MIT Press, 2016." kitabı kullanılarak hazırlanmıştır.

Genel bilgiler

► Değerlendirme

- Ara sınav: 25%
- Ödevler: 15%
- Final projesi: 30%
- Final sınavı: 30%

► Ders kitabı

- I. Goodfellow, Y. Bengio and A. Courville, Deep Learning, MIT Press, 2016.

► Diğer kaynaklar

- S. Russell, and N. Norvig, Artificial Intelligence: A Modern Approach, Prentice Hall, 2003.
- T. Hastie, R. Tibshirani and J. Friedman, The Elements of Statistical Learning, Springer, 2001.
- K. P. Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.

E-Posta : akcayol@gazi.edu.tr, Web : <http://w3.gazi.edu.tr/~akcayol>

Genel bilgiler

Araştırma ödevleri

- ▶ Haftalık konu ile ilgili uygulama içeren bir makale incelenerek detaylı rapor hazırlanacaktır.
- ▶ İncelenen makalede ilgili **yöntemin, algoritmanın, yaklaşımın kullanılmasının gerekçeleri** ve **elde edilen sonuçlar** değerlendirilecektir.
- ▶ İncelenen makale **son 3 yılda yayınlanmış** olacaktır.
- ▶ İncelenen makale **SCI/E tarafından taranan bir dergide** yayınlanmış olacaktır.
- ▶ **Haftalık ödev içeriği:**
 - ▶ İncelenen makalenin tam metni
 - ▶ SCI/E tarafından tarandığını gösterir belge (Thomson Reuters)
 - ▶ Hazırlanacak rapor (Kapak sayfası, İçindekiler, Özet, Materyal/Metot, Sonuçlar, Yorum)

Genel bilgiler

Final projesi

- ▶ Derste anlatılan bir **yöntemin/algoritmanın/yaklaşımın bir alana uygulamasını içerecektir.**
- ▶ Geliştirilecek uygulamanın **algoritma kısmında hazır araç, fonksiyon** veya **kütüphane kullanılmayacaktır.**
- ▶ Hazırlanan projenin; **program kodları, veritabanı ve kütüphane** gibi diğer dokümanları **CD ile, final proje raporu çıktı** olarak teslim edilecektir.
- ▶ **Final projesi içeriği:**
 - ▶ Uygulama alanı hakkında bilgi (seçilme gerekçesi, daha önce ilgili alanda yapılan uygulamalar, ilgili alanın önemi)
 - ▶ Uygulanacak yöntem/algoritma/yaklaşımın seçim gerekçesi (literatürde uygulanan yöntemlerin karşılaştırmalı analizi)
 - ▶ Geliştirilecek uygulamanın sonuçlarının karşılaştırmalı analizi

Ders kapsamı

- Derin öğrenmeye giriş
- Olasılık, dağılımlar ve bilgi teorisi
- Makine öğrenmesinin temelleri
- Yapay sinir ağları
- Deep feedforward networks
- Convolutional neural networks ve uygulamaları
- Recurrent neural networks ve uygulamaları
- Autoencoders ve uygulamaları
- Restricted Boltzmann makinesi ve uygulamaları
- Deep belief networks ve uygulamaları

5

İçerik

- Yapay zeka
- Makine öğrenmesi
- Yapay sinir ağları
- Derin öğrenme

6

Yapay zeka

- Artificial Intelligence (AI), **zeki insan davranışını makinelerin taklit etmesini** amaçlayan çalışmaları kapsar.
- İlk defa McCarthy tarafından 1956'da kullanıldı.
- AI, **iki boyutta** dört farklı şekilde **tanımlanmaktadır**.

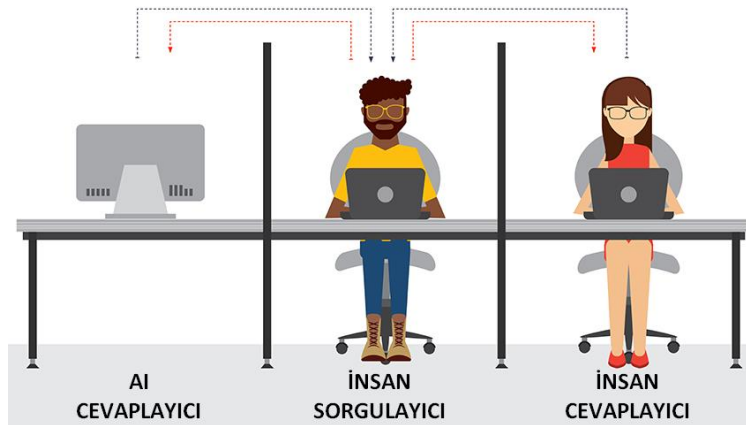
		BAŞARI ÖLÇÜTÜ	
		İNSAN ZEKASINA GÖRE	RASYONELLİĞE GÖRE
TANIM BOYUTU	DÜŞÜNCE	İnsan gibi düşünen sistemler (Bellman, 1978) Cognitive science	Rasyonel düşünen sistemler (Chamlak and McDermott, 1985) Mantık kuralları
	DAVRANIŞ	İnsan gibi davranan sistemler (Kurzweil, 1990) Turing testi	Rasyonel davranan sistemler (Schalkoff, 1990) Rasyonel ajanlar

7

Yapay zeka

İnsan gibi davranan sistemler

- Turing testinde başarılı olan makinelerdir.



8

Yapay zeka

Rasyonel davranan sistemler (rasyonel ajan)

- ▶ **Rasyonel davranış:** Girişlere göre **en doğruyu yapmaktır.**
- ▶ **Doğru:** Verilen bilgiye göre beklenen **faydayı maksimum yapar.**
- ▶ Rasyonel sistemlerin **insan düşüncesini tekrarlaması** veya **aynı kararı vermesi beklenmez.**
- ▶ Bir rasyonel ajan algı geçmişini yeni girişe göre davranışa dönüştüren bir fonksiyondur.

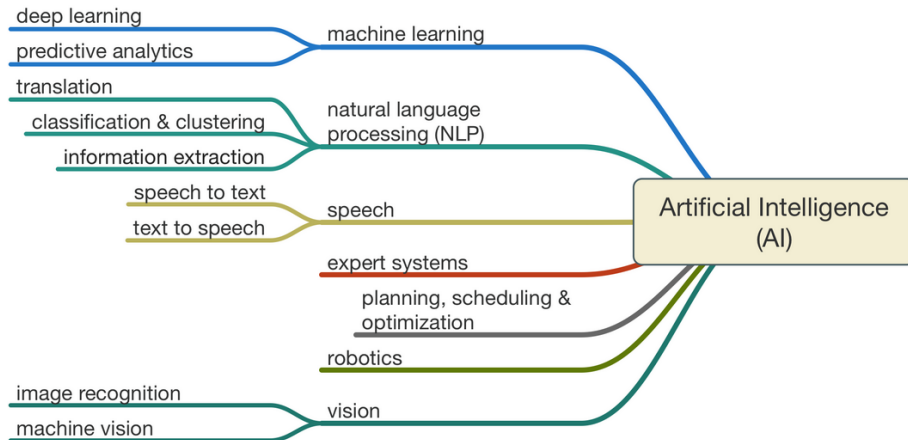
$$D = f(A^*, A)$$

- ▶ D davranış, A^* algı geçmiş, A yeni algı, f fonksiyondur.
- ▶ Rasyonel ajan, **çevresini algılar** (sensörler vb.) ve **en başarılı davranışı belirler.**

9

Yapay zeka

- ▶ Yapay zekanın alanları ve aralarındaki ilişkiler.



10

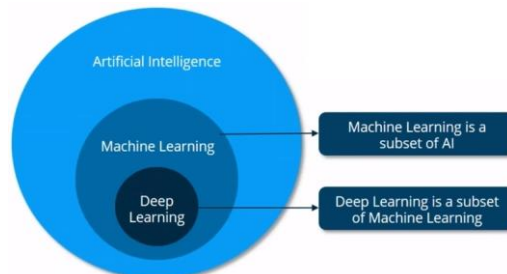
İçerik

- ▶ Yapay zeka
- ▶ **Makine öğrenmesi**
- ▶ Yapay sinir ağları
- ▶ Derin öğrenme

11

Makine öğrenmesi

- ▶ **AI**, insan davranışını makinelere kazandırmayı amaçlar.
- ▶ **Makine öğrenmesi**, yapay zekanın bir alt sınıfıdır ve AI algoritmalarını makinelere uygular.
- ▶ **Artificial Neural Network (ANN)**, makine öğrenmesinin bir alt sınıfıdır ve sinir hücrelerinin çalışmasını taklit eder.
- ▶ **Deep learning**, ANN'in bir alt sınıfıdır ve insan beyninin, algılama ve karar verme becerisini sinir ağları kullanarak taklit etmeyi amaçlar.



12

Makine öğrenmesi

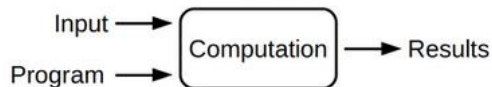
- ▶ **Makine öğrenmesi**, genellikle, **karar veya tahmin oluşturmak** amacıyla öğrenme işlemini gerçekleştirir.
- ▶ Makine öğrenmesi **sınıflandırma (classification)** ve **kümeleme (clustering)** problemlerinde başarıyla uygulanır.
- ▶ **Makine öğrenme yöntemleri** iki türdür:
 - ▶ Supervised learning
 - ▶ Unsupervised learning
- ▶ Supervised learning yönteminde, giriş ile birlikte **çıkış bilgisi** de **sağlanır** (classification).
- ▶ Unsupervised learning yönteminde, **çıkış** etiketleri **sağlanmaz**. Girişler arasındaki ilişki sağlanır (clustering).

13

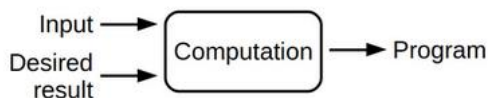
Makine öğrenmesi

- ▶ **Geleneksel programlamada**, giriş ve program sağlanır ve çıkışlar elde edilir.
(**veri + kural kümesi** -> **bilgisayar** -> **sonuçlar**)
- ▶ **Makine öğrenmesinde**, giriş ve istenen çıkışlar sağlanır ve kurallar/program/algoritma elde edilir.
(**veri + sonuçlar** -> **algoritma + bilgisayar** -> **kurallar**)

Traditional programming



Machine learning



14

İçerik

- ▶ Yapay zeka
- ▶ Makine öğrenmesi
- ▶ **Yapay sinir ağları**
- ▶ Derin öğrenme

15

Yapay sinir ağları

- ▶ İnsan beyni, **her görüntüde, seste, tat almada, dokunmada** bir **algı oluşturur**.
- ▶ Beynimiz olmasaydı hepimiz **ilkel organizmalar olurduk** ve sadece **basit refleksler yapabilirdik**.
- ▶ **Bebeğin beyni sadece 1 kilodur**, ancak günümüzdeki çoğu süper bilgisayarların çözemediği problemleri çözer.
- ▶ Doğumdan **birkaç gün sonra, ebeveynlerinin yüzlerini tanıyabilir**, arka planlarından **nesneleri ayırt edebilir** ve hatta **ayrı sesler çıkarabilir**.

16

Yapay sinir ağıları

- ▶ **Bir yıl içinde**, fizik kurallarını algılayacak bir **sezgi geliştirirler**, engellerin arkasındaki **nesneleri takip edebilirler**, **seslerle anlam** ilişkilendirmesi yapabilirler.
- ▶ **Çocukluk çağında**, binlerce kelime ve **karmaşık gramer yapısını anlayabilirler**.
- ▶ Son yıllarda, **insan beyni gibi beyne sahip zeki makineler** geliştirilmeye çalışılmaktadır.

17

Yapay sinir ağıları

- ▶ Bilgisayar programları **aritmetik işlemleri hızlı yapabilir** veya **sıralı komutları hızlı çalıştırabilir**.
- ▶ Ancak, klasik programlar **bir kişinin el yazısını okuma** gibi işlemlerde çok **başarılı değildir**.



18

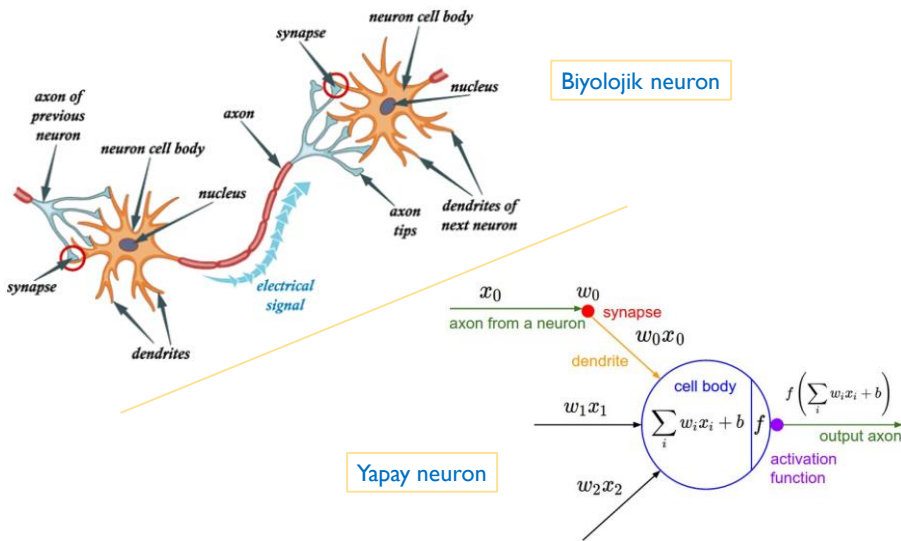
Yapay sinir ağıları

- ▶ Yapay sinir ağıları, **insan beyninin yapısını ve fonksiyonunu taklit eder.**
- ▶ Sınıflandırıcı $y = f(x)$, x **girişi ile y sınıfını eşleştirir.**
- ▶ Bir feedforward ANN, $y = f(x, w)$, w parametrelerinin değerlerini öğrenerek daha iyi tahmin yapan bir fonksiyon elde eder.
- ▶ **Öğrenme supervised** veya **unsupervised** olabilir.
- ▶ Yapay sinir ağıları, **metin, ses, görüntü, video** gibi **yapılandırılmamış veriyi alır** ve neuronlardan oluşan **katmanlarda işler.**
- ▶ **Neuron'lar** başka neuron'ları aktif yapabilir veya dış ortamda **başka işlemleri başlatabilir.**

19

Yapay sinir ağıları

- ▶ Yapay sinir ağı, çok sayıda **işlem birimine (neuron)** sahiptir.

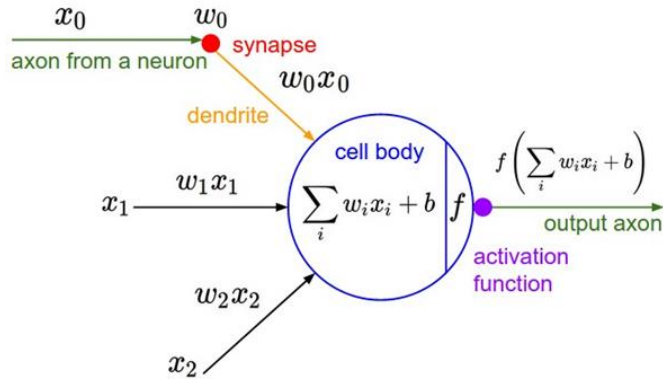


20

Yapay sinir ağıları

Perceptron

- Bir **perceptron** tek katmanlı bir **neural network**'tür.



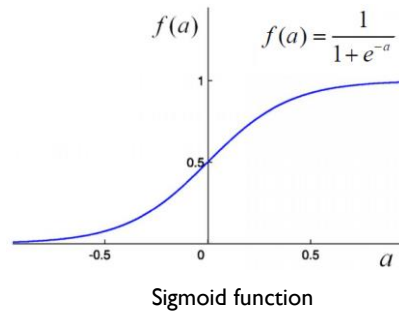
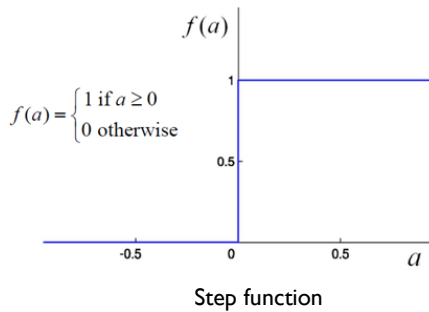
- $x = (x_0, x_1, x_2)$ giriş vektörü, $w = (w_0, w_1, w_2)$ ağırlık vektörü ve b bias, $y = f(x_0 w_0 + x_1 w_1 + x_2 w_2 + b)$ olur.

21

Yapay sinir ağıları

Perceptron

- Transfer fonksiyonuna göre çıkış değeri değişir.

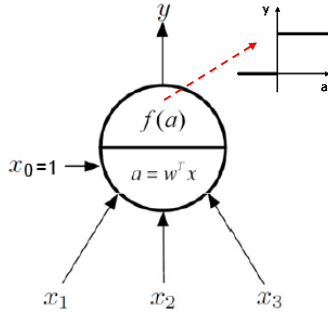


22

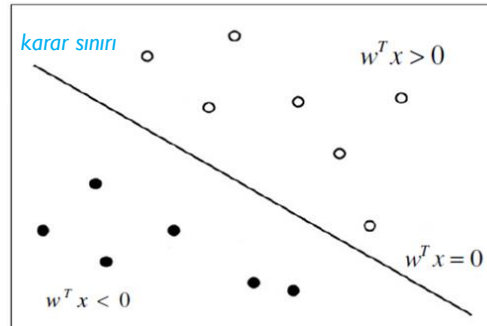
Yapay sinir ağıları

Perceptron limiti

- Tek neuron bir doğru ile çözüm uzayını ayırır.



$$y = f(a) = \begin{cases} 1 & \text{if } a \geq 0 \\ 0 & \text{otherwise} \end{cases} \Leftrightarrow y(x) = \begin{cases} 1 & \text{if } w^T x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$



$$x = (1, x_1, x_2, x_3)$$

$$w = (w_0, w_1, w_2, w_3)$$

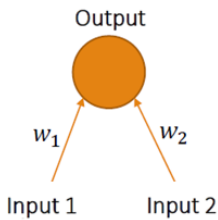
$$a = w^T x$$

23

Yapay sinir ağıları

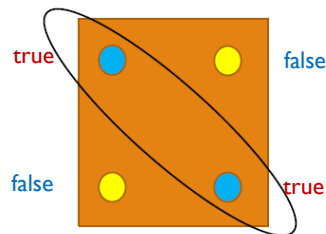
Perceptron limiti

- Tek perceptron XOR problemini çözemez (θ eşik değerdir).



Input 1	Input 2	Output
1	1	0
1	0	1
0	1	1
0	0	0

- $0 w_1 + 0 w_2 < \theta \rightarrow 0 < \theta$ Tutarlı!!!
- $0 w_1 + 1 w_2 > \theta \rightarrow w_2 > \theta$
- $1 w_1 + 0 w_2 > \theta \rightarrow w_1 > \theta$
- $1 w_1 + 1 w_2 < \theta \rightarrow w_1 + w_2 < \theta$

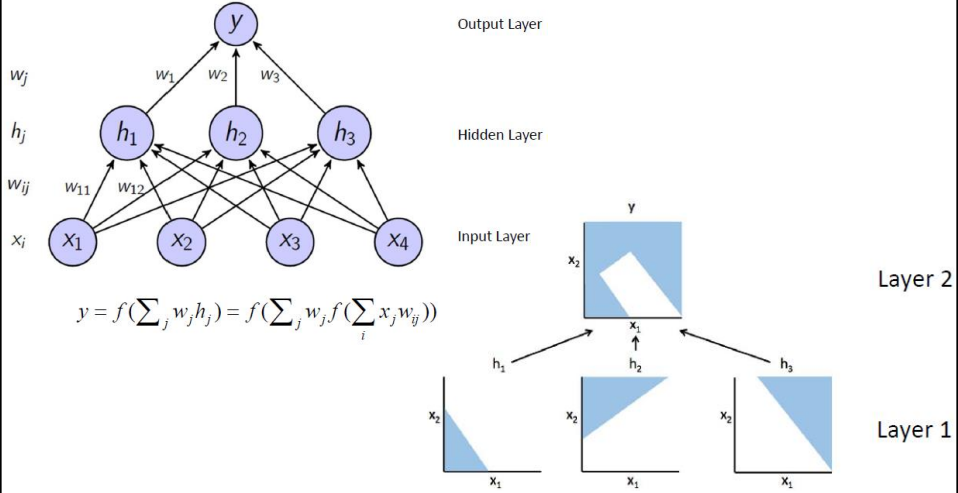


24

Yapay sinir ağıları

Çok katmanlı neural network

- Çok katmanlı ANN daha karmaşık problemleri çözebilir.

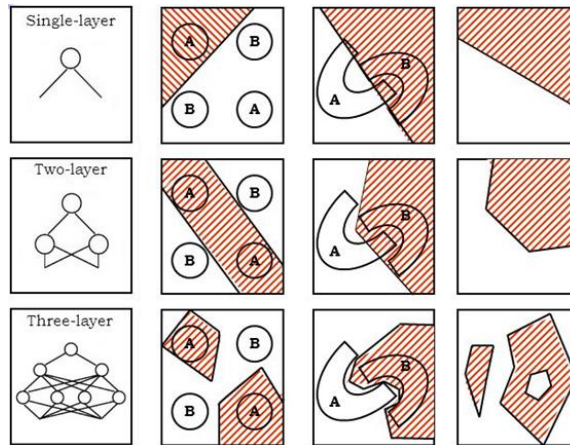


25

Yapay sinir ağıları

Çok katmanlı neural network

- Katman ve neuron sayısı arttıkça çok daha karmaşık problemleri çözebilir.

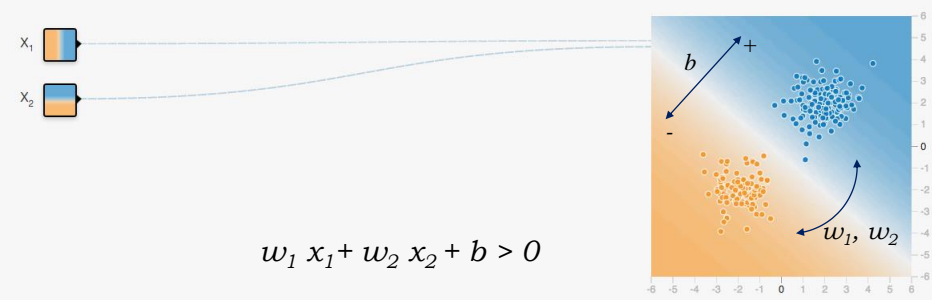


26

Yapay sinir ağıları

Çok katmanlı neural network

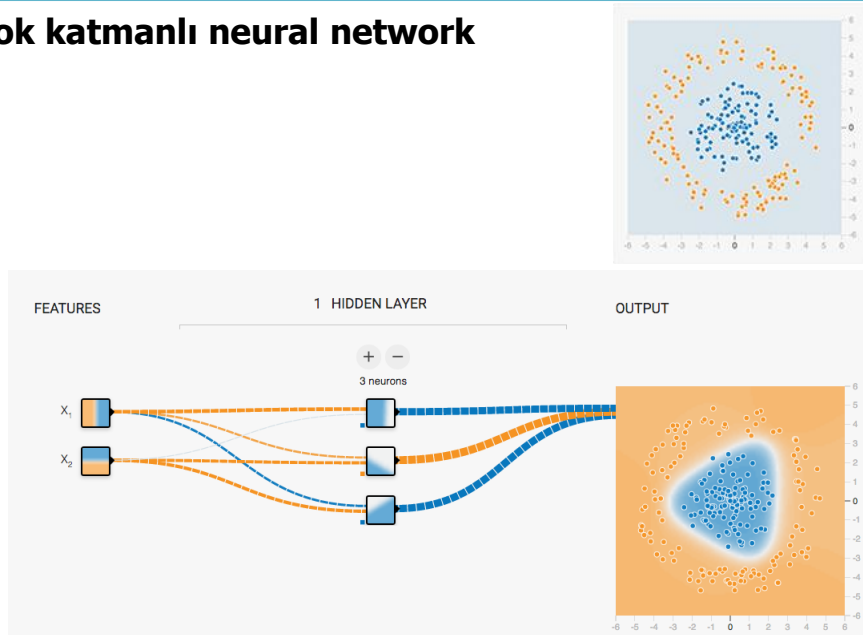
- ▶ w_1, w_2, b parametredir.
- ▶ w_1, w_2 doğruyu döndürür, b doğrunun yerini değiştirir.
- ▶ w_1, w_2 ve b için en uygun değerlerin belirlenmesi gereklidir.



27

Yapay sinir ağıları

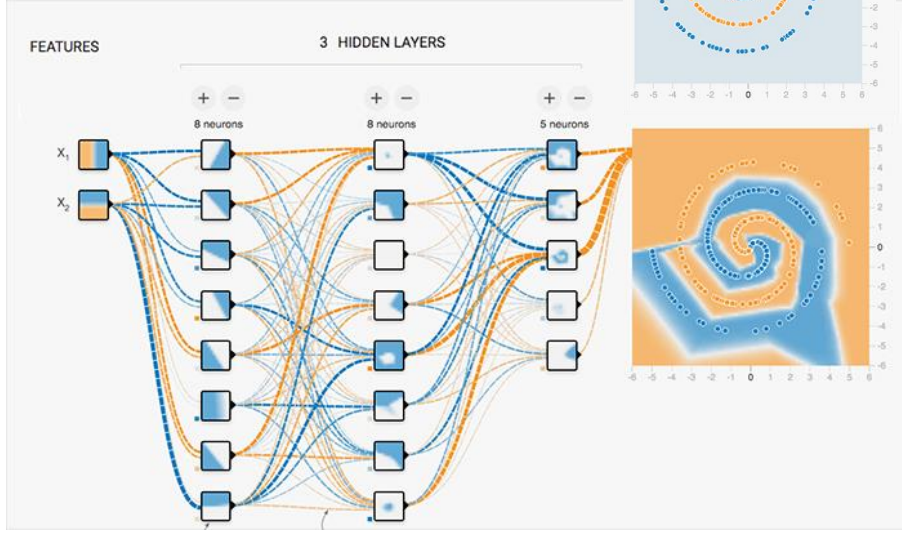
Çok katmanlı neural network



28

Yapay sinir ağıları

Çok katmanlı neural network



29

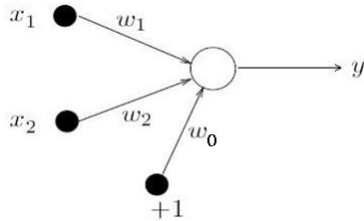
Yapay sinir ağıları

Perceptron öğrenme

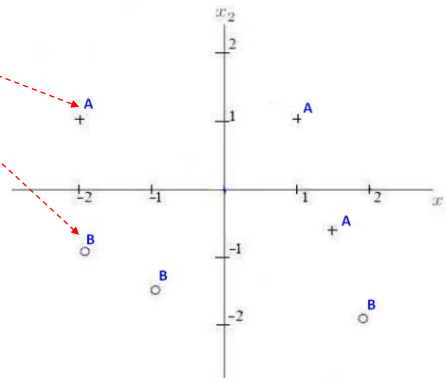
- ▶ Perceptron öğrenmesinde, w_1 , w_2 ve θ değeri belirlenir.
- ▶ η **öğrenme hızı**, $d(n)$ istenen **çıkış**.
- ▶ Yanlış sınıflandırmada ağırlık değiştirilir.

$$y = f(w_1x_1 + w_2x_2 + w_0)$$

$$d(n) = \begin{cases} +1 & \text{if } x(n) \in \text{set A} \\ -1 & \text{if } x(n) \in \text{set B} \end{cases}$$



$$w_i = w_i + \eta (d(n) - y(n)) x_i(n)$$



30

Yapay sinir ağıları

Perceptron öğrenme

$w^T x < 0$ ise \ominus
 $w^T x > 0$ ise \oplus

$$w_i = w_i - \eta (d(n) - y(n)) x_i(n)$$

Başlangıç değerleri

$$\eta = 0.2$$

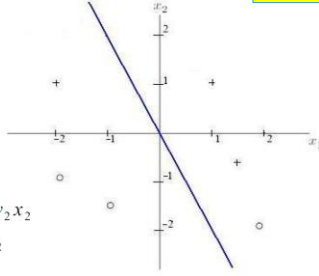
$$w = \begin{pmatrix} 0 \\ 1 \\ 0.5 \end{pmatrix}$$

$$x_0 = +1$$

$$0 = w_0 + w_1 x_1 + w_2 x_2$$

$$= 0 + x_1 + 0.5x_2$$

$$\Rightarrow x_2 = -2x_1$$



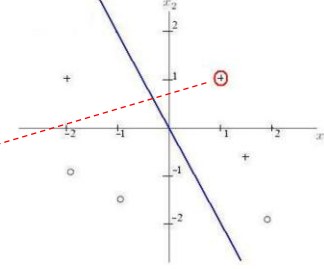
$$\eta = 0.2$$

$$w = \begin{pmatrix} 0 \\ 1 \\ 0.5 \end{pmatrix}$$

$$x_1 = 1, x_2 = 1$$

$$w^T x > 0$$

Doğru sınıflandırma



$$\eta = 0.2$$

$$w = \begin{pmatrix} 0 \\ 1 \\ 0.5 \end{pmatrix}$$

$$x_1 = 2, x_2 = -2$$

$$w_0 = w_0 - 0.2 * 1$$

$$w_1 = w_1 - 0.2 * 2$$

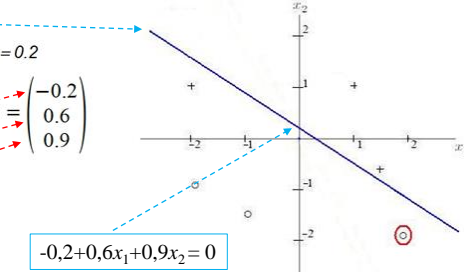
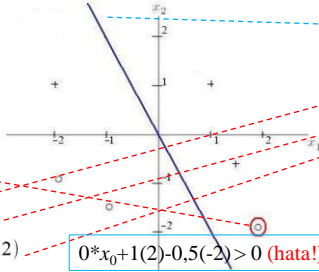
$$w_2 = w_2 - 0.2 * (-2)$$

$$0 * x_0 + 1(2) - 0.5(-2) > 0 \text{ (hata!)}$$

$$\eta = 0.2$$

$$w = \begin{pmatrix} -0.2 \\ 0.6 \\ 0.9 \end{pmatrix}$$

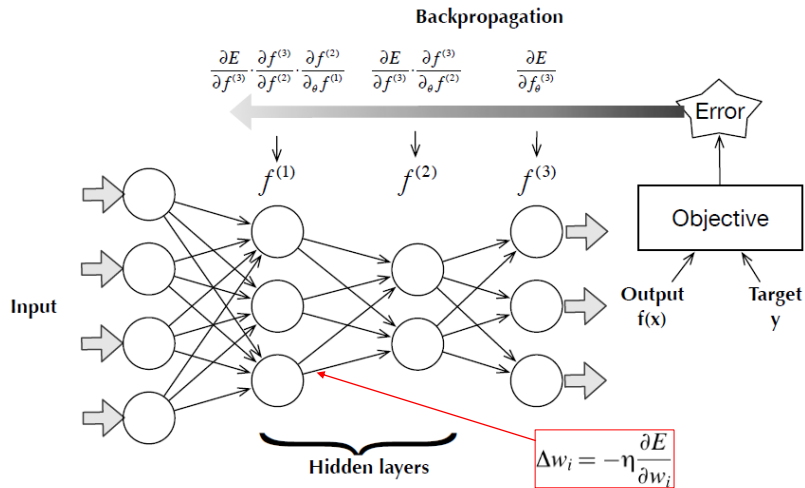
$$-0.2 + 0.6x_1 + 0.9x_2 = 0$$



Yapay sinir ağıları

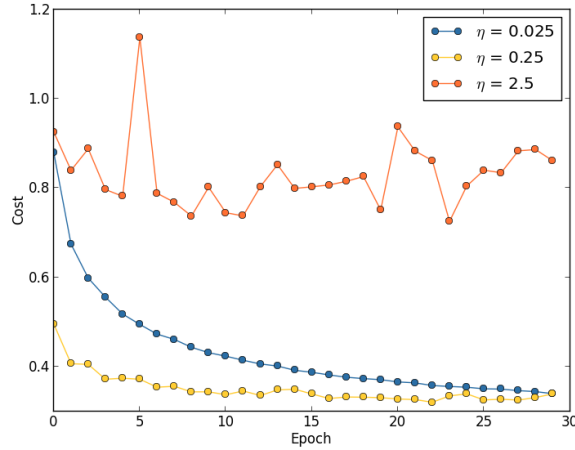
Çok katmanlı feedforward ANN öğrenme

► Çok katmanlı ANN'de backpropagation yöntemi kullanılır.



Yapay sinir ağıları

- Öğrenme oranının uygun seçilmesi gerekir.
- Yüksek seçilirse kararsızlık ortaya çıkar, düşük seçilirse öğrenme uzun sürer.



33

İçerik

- Yapay zeka
- Makine öğrenmesi
- Yapay sinir ağıları
- **Derin öğrenme**

34

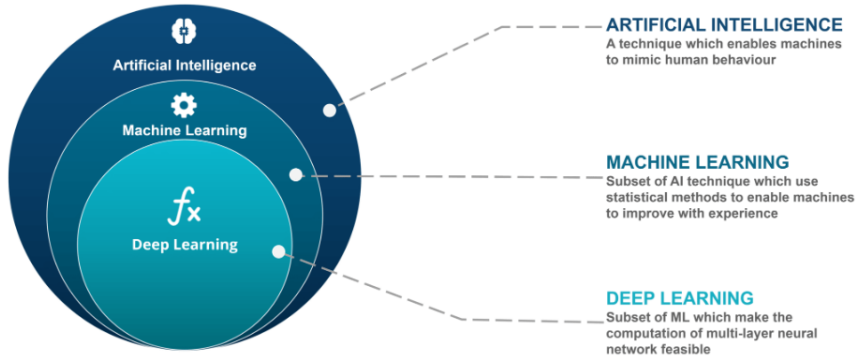
Derin öğrenme

- ▶ Deep learning, son yıllarda klasik yapay sinir ağlarının yerini almıştır.
- ▶ Deep learning, **makine öğrenme tekniklerinin** ve mimarilerinin geniş **bir sınıfını ifade eder.**
- ▶ **Shallow learning**, bir veya iki katman kullanarak yapılan öğrenmeyi ifade eder.
- ▶ **Deep learning**, çok sayıda katmanla öğrenmeyi ifade eder.
- ▶ Deep learning, işlem birimi olan **neuron'larda nonlinear fonksiyon kullanır.**

35

Derin öğrenme

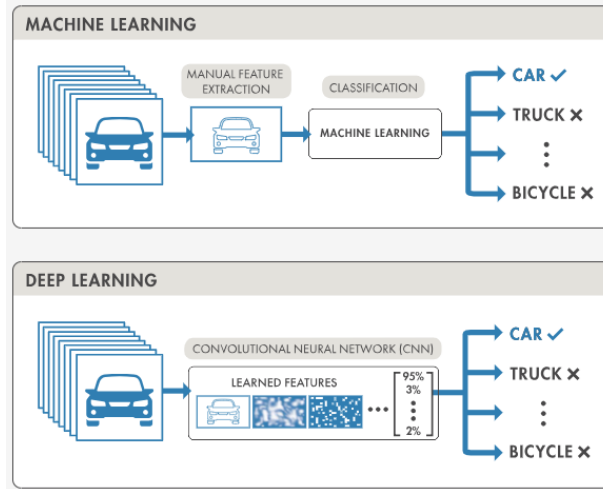
- ▶ **Makine öğrenmesinde ve derin öğrenmede**, makinelere geçmiş verileri kullanarak tahmin yapma ve öğrenme yeteneği kazandırılır
- ▶ **Makine öğrenmesi**, genellikle **istatistiksel metotlar** yaygın kullanılır; **derin öğrenme** ise çok katmanlı **sinir ağları** ile hesaplama yapar.



36

Derin öğrenme

- **Makine öğrenmesinde**, genellikle öznitelik çıkarımı manuel yapılır, **derin öğrenmede** öznitelikler de otomatik olarak öğrenilebilir.



37

Derin öğrenme

- Yaygın kullanılan derin öğrenme modelleri:
 - Deep feedforward networks
 - Convolutional neural networks
 - Recurrent neural networks
 - Autoencoders
 - Deep belief networks

38

Ödev

- Derin öğrenmenin sağlık alanında uygulamasını içeren SCI/E dergilerinde yayınlanmış bir makale hakkında ödev hazırlayınız.