

Derin Öğrenme Deep Learning

Hazırlayan: M. Ali Akcayol
Gazi Üniversitesi
Bilgisayar Mühendisliği Bölümü

Bu dersin sunumları, "Simon Haykin, Neural Networks and Learning Machines, Prentice Hall, 2016." kitabı kullanılarak hazırlanmıştır.

İçerik

- ▶ Yapay sinir ağı nedir?
- ▶ Yapay sinir ağının özellikleri
- ▶ İnsan beyni
- ▶ Neuron modeli
- ▶ Yapay sinir ağı mimarileri
- ▶ Öğrenme işlemleri

Yapay sinir ađı nedir?

- ▶ **Biyolojik sinir sisteminden esinlenerek** ortaya çıkmıştır.
- ▶ İnsan beyninin, **öđrenme, eski bilgiye dayalı tahmin etme, eksik bilgiyi tamamlama** gibi yeteneklerini makinelere kazandırmayı amaçlar.
- ▶ **Çok sayıda işlem ünitesinden (processing elements, units, nodes, neurons) oluşur.**
- ▶ Birçok farklı alanda başarıyla uygulanmıştır:
 - ▶ İşaret tanıma
 - ▶ El yazısı tanıma
 - ▶ İleriye dönük tahminde bulunma
 - ▶ Adaptif kontrol
 - ▶ Gürültülü veya eksik bilgileri tanıma
 - ▶ ...

3

Yapay sinir ađı nedir?

- ▶ **Beyin, çok karmaşık** yapıdadır, **dođrusal olmayan** ve **paralel işlem** yapar.
- ▶ Beyin, **neuron**'lardan (**processing units**) oluşur ve bazı problemlerde günümüzdeki en hızlı bilgisayarlardan daha hızlı hesaplama kapasitesine sahiptir.
- ▶ Yapay sinir ađı (YSA), **beynin bazı özel işlevlerini modellemek için tasarlanan makinedir.**
- ▶ YSA, iyi performans sağlamak için çok sayıda basit işlem birimi (neuron) arasında bağlantı yapar.
- ▶ YSA ile beyin iki açıdan benzerdir:
 - ▶ **Bilgi**, ađ tarafından öđrenme sürecinde **çevreden edinilir.**
 - ▶ Neuronlar arasındaki bağlantının gücü (**synaptic weights**) kullanılarak/deđiştirilerek **edinilen bilgi saklanır.**
- ▶ Öđrenme sürecinde kullanılan **öđrenme algoritması ile ađırlıklar/bađlantı gücü deđiştirilir.**

4

Yapay sinir ađı nedir?

- ▶ Walter Pitts ve Warren McCulloch tarafından **1943 yılında** biyolojik nöronların **ilk yapay nöron modeli** oluşturulmuştur.
- ▶ Bütün **mantıksal işlemler** farklı eşik değerleri ve farklı ağırlıklarla **gerçekleştirilmiştir**.
- ▶ Donald Hebb tarafından 1949 yılında **Hebb öğrenme** kuralı geliştirilmiştir.
- ▶ Hebb kuralına göre, **iki nöron aktifse** ikisinin **aralarındaki bağlantı güçlendirilir**.
- ▶ Rosenblatt tarafından 1958 yılında **ilk perceptron** ve **öğrenme metodu geliştirildi**.
- ▶ Widrow ve Hoff tarafından 1960 yılında **Adaptive Linear Element (ADALINE)** geliştirildi.
- ▶ Widrow ve Hoff, türevlenebilir fonksiyonlarla **gradient descent** tabanlı öğrenme kuralını geliştirmiştir.

5

Yapay sinir ađı nedir?

- ▶ Rosenblatt 1961 yılında **backpropagation öğrenme** şemasını önerdi, ancak **ađın eğitiminde başarılı olamadı**.
- ▶ Minsky ve Papert 1969 yılında **perceptron'un** bazı basit mantıksal işlemlerde **yetersiz olduğunu göstermiştir**.
- ▶ Bir perceptron ile **XOR** probleminin çözülemediđi gösterilmiştir.
- ▶ **Çok katmanlı ađ ile bu tür problemlerin çözülebileceđi düşünöldü**, ancak nasıl eğitilebileceđi konusunda çözüm bulunamadı.
- ▶ 1980'li yıllarda bu tür problemler çok katmanlı ve farklı ađ yapıları kullanılarak çözüldü.
- ▶ Günümüzde, **YSA diđer yapay zeka teknikleriyle** (fuzzy logic, genetik algoritma) **birlikte** kullanılarak çok daha **etkin çözümler ortaya koymaktadır**.

6

İçerik

- ▶ Yapay sinir ağı nedir?
- ▶ **Yapay sinir ağının özellikleri**
- ▶ İnsan beyni
- ▶ Neuron modeli
- ▶ Yapay sinir ağı mimarileri
- ▶ Öğrenme işlemleri

7

Yapay sinir ağının özellikleri

- ▶ **YSA**, birbirine çok sayıda bağlantıya sahip olan çok sayıda neuron ile **öğrenme** ve **genelleme** özelliğine sahiptir.
- ▶ **Genelleme**, sinir ağlarının eğitim sırasında **görmediği girişler** için **makul/doğru çıkışlar üretmesidir**.
- ▶ YSA, karmaşık ve zor problemler için iyi düzeyde çözümler üretebilir.
- ▶ YSA aşağıdaki önemli özelliklere ve yeteneklere sahiptir:
 - ▶ Nonlineer
 - ▶ Giriş-çıkış eşleştirme
 - ▶ Adaptif
 - ▶ Bağlamsal bilgi
 - ▶ Hata toleransı
 - ▶ Nörobiyolojik analogi

8

Yapay sinir ağının özellikleri

Nonlinear

- ▶ Bir **yapay neuron**, **linear** veya **nonlinear** olabilir.
- ▶ **Girişin yüksek nonlinear olduğu durumlarda** (ses sinyali) YSA'nın nonlinear olma özelliği çok önemlidir.

Giriş-çıkış eşleştirme

- ▶ Popüler öğrenme yaklaşımı olan **denetimli öğrenmede** (**supervised learning**) ağırlıklar giriş örneklerine göre değiştirilir.
- ▶ YSA, öğrenme süreci sonunda giriş-çıkış eşleştirmesi için gereken ağırlık değerlerini oluşturur.

Adaptif

- ▶ YSA, **ağırlıklarını** çevresindeki **değişimlere adapte edebilir**.
- ▶ Bir problem için eğitilen sinir ağı, **problemin değişen şartları için yeniden eğitilebilir**.

9

Yapay sinir ağının özellikleri

Bağlamsal bilgi

- ▶ Her neuron, YSA'da **diğer neuronların aktivitesinden** belirli oranda **etkilenir**.

Hata toleransı

- ▶ YSA, **hata toleransına sahiptir**.
- ▶ Olumsuz şartlarda (belirli oranda giriş eksik) performansı düşer (**graceful degradation**).

Nörobiyolojik analogi

- ▶ YSA beyin taklit edilerek tasarlanır.
- ▶ **Beyin, hata toleranslı paralel işlemin fiziksel olabilirliğini**, aynı zamanda **hızlı** ve **güçlü** olduğunu ispatlar.

10

İçerik

- ▶ Yapay sinir ağı nedir?
- ▶ Yapay sinir ağının özellikleri
- ▶ **İnsan beyni**
- ▶ Neuron modeli
- ▶ Yapay sinir ağı mimarileri
- ▶ Öğrenme işlemleri

11

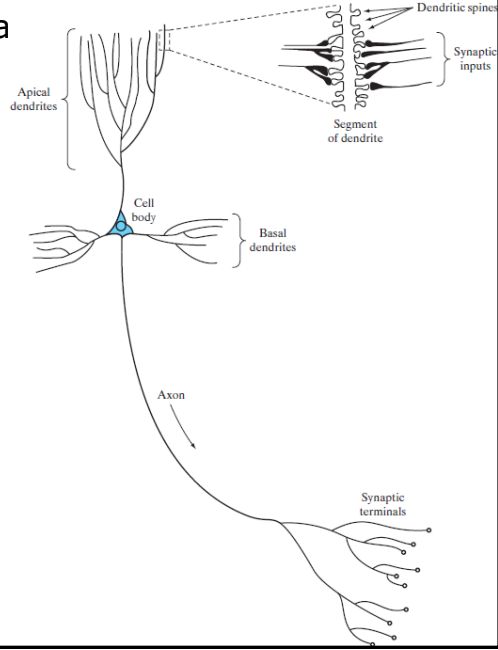
İnsan beyni

- ▶ İnsan beyinde **10 milyar neuron** ve **60 trilyon bağlantı** olduğu tahmin edilmektedir.
- ▶ **Sinapslar** veya **sinir uçları**, **neuronlar arasında etkileşimi sağlayan** temel bileşenlerdir.
- ▶ **Beyin plastisitesi**, **beynin bağlantılarını düzenleme ya da yeni bağlantılar kurma yetisidir.**
- ▶ **Plastisite**, sinir sisteminin çevresine adapte olarak gelişmesine izin verir.
- ▶ **Plastisite ile;**
 - ▶ Neuronlar arasında **yeni sinaptik bağlantılar oluşturulur.**
 - ▶ Varolan sinaptik **bağlantılar değiştirilir.**
- ▶ Beyinde farklı kısımlardaki neuronlar farklı boyutlarda ve şekillerdedir.

12

İnsan beyni

- ▶ Bir neuron, **10.000** veya daha fazla **bağlantı alabilir** ve **binlerce** hedef hücreye **çıkış üretebilir**.
- ▶ Neuronlar arasındaki bağlantıyı **axon**'lar sağlar.
- ▶ Neuronlar **dendrite**'ler ile giriş alma işlemini yapar.

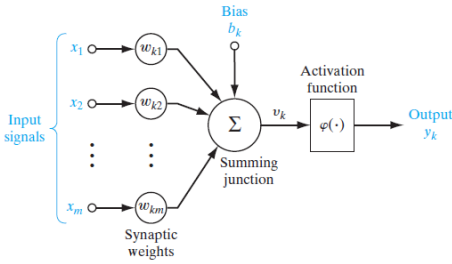


İçerik

- ▶ Yapay sinir ağı nedir?
- ▶ Yapay sinir ağının özellikleri
- ▶ İnsan beyni
- ▶ **Neuron modeli**
- ▶ Yapay sinir ağı mimarileri
- ▶ Öğrenme işlemleri

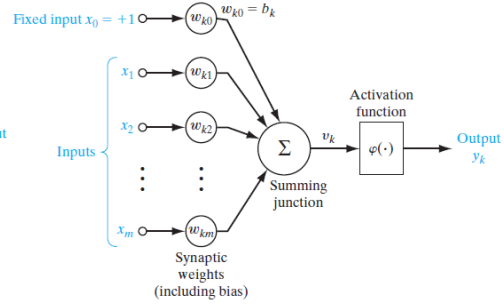
Neuron modeli

- Bir **neuron**, sinir ağının **temel işlem birimidir**.



$$u_k = \sum_{j=1}^m w_{kj} x_j$$

$$v_k = u_k + b_k$$



$$v_k = \sum_{j=0}^m w_{kj} x_j$$

$$y_k = \varphi(v_k)$$

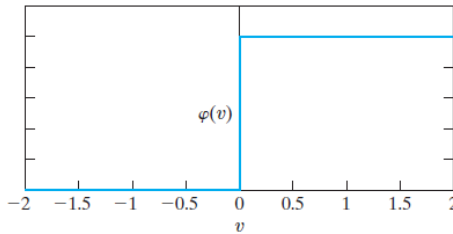
- Bir grup **sinaps** vardır ve **belirli bir ağırlığa** sahiptir.
- **Toplayıcı**, girişlerden gelen sinyalleri doğrusal olarak birleştirir.
- **Aktivasyon fonksiyonu**, neuron'un çıkış genliğini belirler.

15

Neuron modeli

Aktivasyon fonksiyonları

- **Threshold fonksiyonu**, çıkışı bir eşik değere göre belirler.

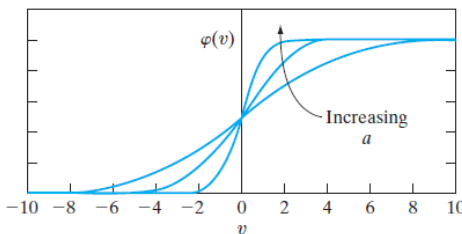


$$\varphi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{if } v < 0 \end{cases}$$

$$y_k = \begin{cases} 1 & \text{if } v_k \geq 0 \\ 0 & \text{if } v_k < 0 \end{cases}$$

$$v_k = \sum_{j=1}^m w_{kj} x_j + b_k$$

- **Sigmoid**, en yaygın kullanılan aktivasyon fonksiyonudur



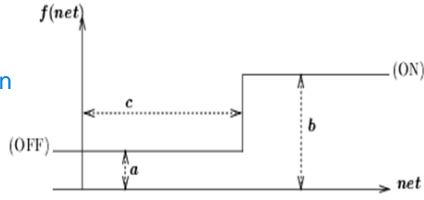
$$\varphi(v) = \frac{1}{1 + \exp(-av)}$$

16

Neuron modeli

Aktivasyon fonksiyonları

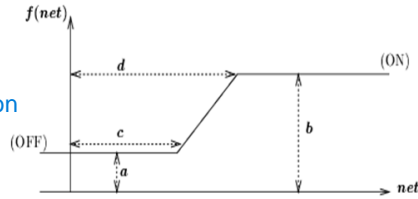
Step function



$$b, \quad \text{net} > c$$

$$a, \quad \text{diğer}$$

Ramp function

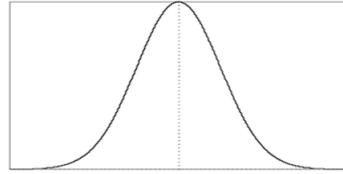


$$b, \quad \text{net} > d$$

$$a, \quad \text{net} < c$$

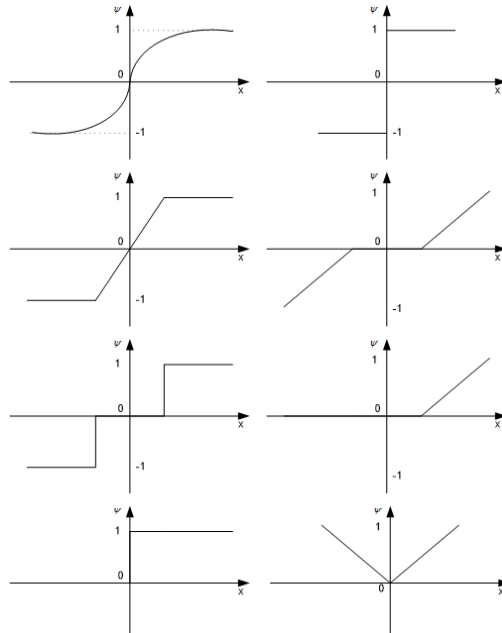
$$a + (net - c)(b - a) / (d - c), \quad \text{diğer}$$

Gaussian function



$$f(\text{net}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{\text{net} - \mu}{\sigma}\right)^2\right]$$

Neuron modeli



İçerik

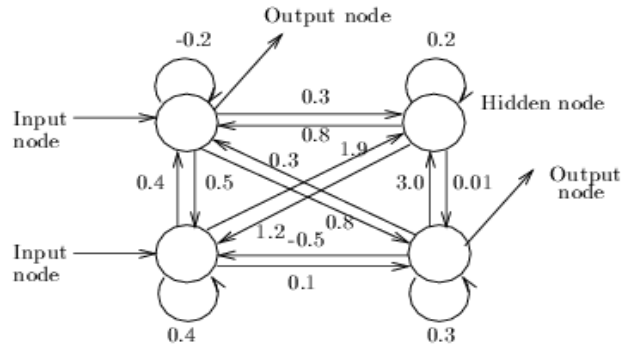
- ▶ Yapay sinir ağı nedir?
- ▶ Yapay sinir ağının özellikleri
- ▶ İnsan beyni
- ▶ Neuron modeli
- ▶ **Yapay sinir ağı mimarileri**
- ▶ Öğrenme işlemleri

19

Yapay sinir ağı mimarileri

Komple bağlı ağlar (fully connected networks)

- ▶ **Her node** diğer **tüm node'lara bağlıdır.**
- ▶ Bağlantılar pozitif, negatif veya 0 olabilir.

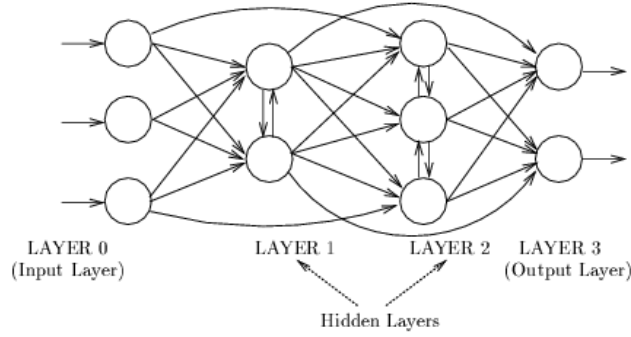


20

Yapay sinir ağı mimarileri

Katmanlı ağlar (layered networks)

- ▶ Node'lar katman ismi verilen alt gruplar halindedir.
- ▶ İleri katmandan **geriye doğru bağlantı olamaz.**

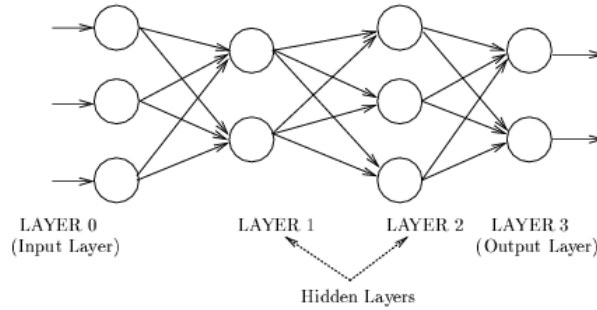


21

Yapay sinir ağı mimarileri

Katmanlı ileri beslemeli ağlar (layered feedforward networks)

- ▶ Bağlantılar, i .katmandaki neuron ile $i+1$.katmandaki neuronlar arasında yapılabilir.
- ▶ En yaygın kullanılan ağ yapısıdır.

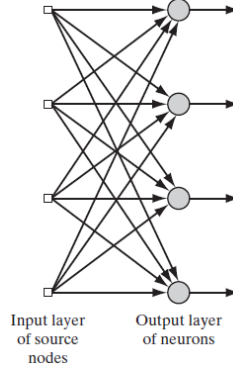


22

Yapay sinir ağı mimarileri

Tek katmanlı ileri beslemeli ağlar

- ▶ YSA'daki neuronların bağlantı şekli **öğrenme algoritmasıyla doğrudan ilişkilidir**.
- ▶ Ağda neuronlar **katman** halinde bağlantıya sahiptir.
- ▶ Ağda **girişlerden çıkışlara doğru** geçiş yapılır (**feedforward**).
- ▶ Girişler ve çıkışlar arasında tek katman vardır.

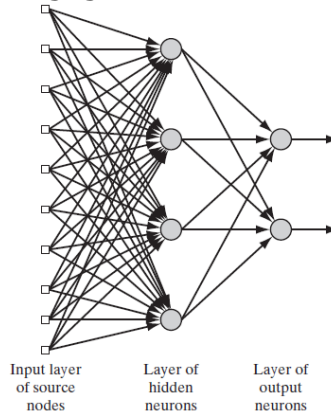


23

Yapay sinir ağı mimarileri

Çok katmanlı ileri beslemeli ağlar

- ▶ Bir veya daha fazla **gizli katman (hidden layer)** vardır.
- ▶ **Gizli katman sayısı artırılarak işlem kapasitesi artırılır.**
- ▶ Üç veya **daha fazla boyutlu** tasarım ile işlem kapasitesi artırılır.
- ▶ Şekilde, 10-4-2 sinir ağı görülmektedir.

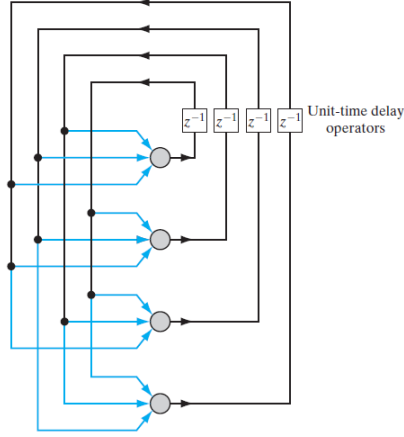


24

Yapay sinir ağı mimarileri

Recurrent networks

- ▶ **Recurrent** (tekrarlayan) **neural network (RNN)**, en az bir tane **feedback döngüsüne sahiptir.**
- ▶ Şekilde, **kendi kendisine feedback olmayan** RNN ağı yapısı görülmektedir.

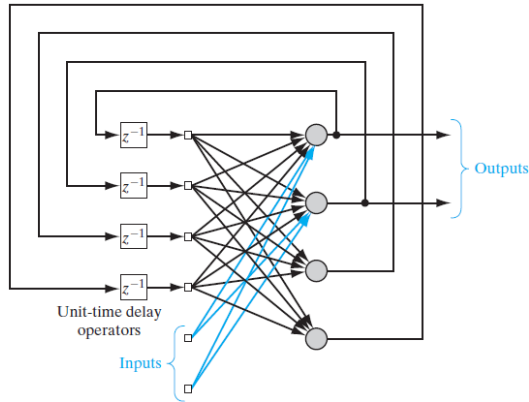


25

Yapay sinir ağı mimarileri

Recurrent networks

- ▶ Şekilde, **kendi kendisine feedback olan** RNN yapısı görülmektedir.



26

İçerik

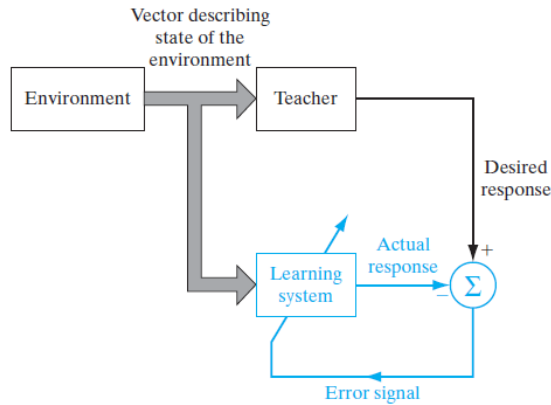
- ▶ Yapay sinir ağı nedir?
- ▶ Yapay sinir ağının özellikleri
- ▶ İnsan beyni
- ▶ Neuron modeli
- ▶ Yapay sinir ağı mimarileri
- ▶ Öğrenme işlemleri

27

Öğrenme işlemleri

Denetimli öğrenme (supervised learning)

- ▶ Denetimli öğrenmede, **giriş ve çıkış** örnekleri birlikte sağlanır.
- ▶ Girişlere göre YSA'dan **beklenen cevap da sağlanır.**
- ▶ YSA parametreleri **hata sinyaline göre ayarlanır.**



28

Öğrenme işlemleri

Denetimli öğrenme (supervised learning)

- ▶ **Toplam hata**, hataların karelerinin toplamı (sum of squared errors) veya hataların karelerinin toplamının ortalaması (mean squared error) ile hesaplanır.

$$SSE = \sum_{p=1}^P \sum_{j=1}^K (d_{p,j} - o_{p,j})^2$$

$$MSE = \frac{1}{P} \sum_{p=1}^P \sum_{j=1}^K (d_{p,j} - o_{p,j})^2$$

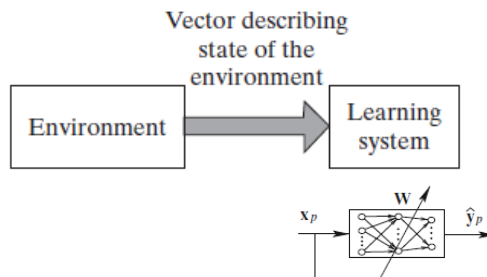
- ▶ **Toplam hata**, belirlenen bir **eşik değerin altına düştüğünde eğitim sonlandırılır.**
- ▶ Eğitim hatası yerine **test hatası minimum olduğunda da eğitim sonlandırılabilir.**

29

Öğrenme işlemleri

Denetimsiz öğrenme (unsupervised learning)

- ▶ Denetimsiz öğrenmede, **sadece girişler** sağlanır.
- ▶ **Giriş değerlerine bağlı verilen ölçüt/kural/ilişki** kullanılarak beklenen cevap oluşturulur.
- ▶ Modeldeki **parametreler giriş değerlerine göre ayarlanır.**



30

Öğrenme işlemleri

Takviyeli öğrenme (reinforcement learning)

- ▶ Takviyeli öğrenmede, **öğrenen sisteme dışarıdan yönlendirme yapılır.**
- ▶ Çıkış değerleri girmek yerine, verilen giriş değerlerine göre **çıkışı değerlendirerek** öğrenmeyi sağlar (genetik algoritma).
- ▶ **Amaç fonksiyonunun değeri minimize/maksimize edilir.**

