

Derin Öğrenme Deep Learning

Hazırlayan: M. Ali Akcayol
Gazi Üniversitesi
Bilgisayar Mühendisliği Bölümü

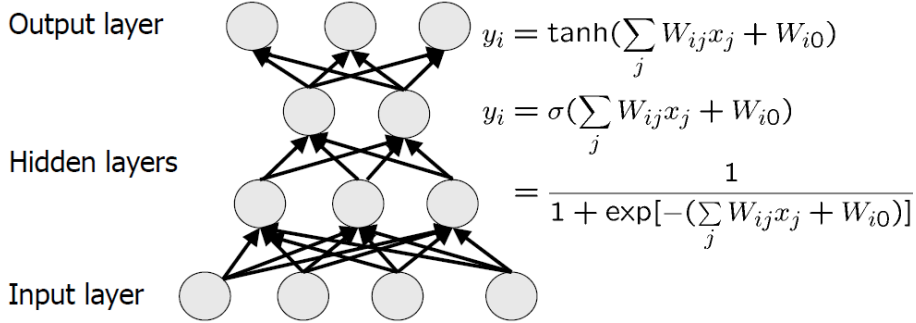
Bu dersin sunumları, "Simon Haykin, Neural Networks and Learning Machines, Prentice Hall, 2016." kitabı kullanılarak hazırlanmıştır.

İçerik

- ▶ Derin mimariler
- ▶ Derin inanç ağları
- ▶ DBN yapısı
- ▶ DBN eğitimi
- ▶ DBN parametreleri
- ▶ DBN uygulamaları

Derin mimariler

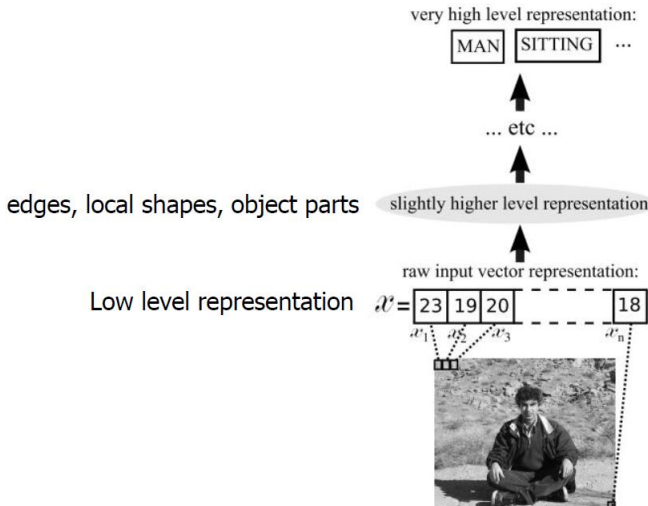
- ▶ **Derin mimariler çok seviyeli non-linear işlem biriminden oluşur.**
- ▶ **Çok sayıda gizli katmana sahip yapay sinir ağları** derin mimariye sahiptir.



3

Derin mimariler

- ▶ **Derin öğrenme yöntemleri özellik hiyerarşisini öğrenir.**
- ▶ **Yüksek düzeyli özellikler** düşük düzeyli özelliklerden **hiyerarşik olarak elde edilir.**



4

Derin mimariler

- ▶ Bazı **karmaşık problemler** shallow neural network yapısıyla **istenen düzeyde çözülememektedir.**
- ▶ **Backpropagation** ile eğitim için **etiketlenmiş veriye ihtiyaç duyulur.**
- ▶ Günümüzdeki **verilerin büyük bölümü etiketlenmiş değildir.**
- ▶ Çok katmanlı ağlarda **backpropagation** ile **öğrenme süresi genellikle çok uzundur.**

5

İçerik

- ▶ Derin mimariler
- ▶ **Derin inanç ağları**
- ▶ DBN yapısı
- ▶ DBN eğitimi
- ▶ DBN parametreleri
- ▶ DBN uygulamaları

6

Derin inanç ağıları

- ▶ **Klasik MLP ayırıcı (discriminative) modeldir, derin inanç ağıları (deep belief networks - DBN) üretici (generative) modeldir.**
- ▶ **DBN, gizli değişkene sahip olan çok sayıda katmana sahiptir.**
- ▶ **Her katman alt katmandaki özelliklerin arasındaki üst düzey ilişkiyi elde eder.**
- ▶ **DBN'lerin en üstteki iki katmanı Restricted Boltzmann Machine yapısına sahiptir ve yönsüz bipartite graftır.**
- ▶ **Alt katmanlar ise yönlü sigmoid belief network yapısına sahiptir.**

7

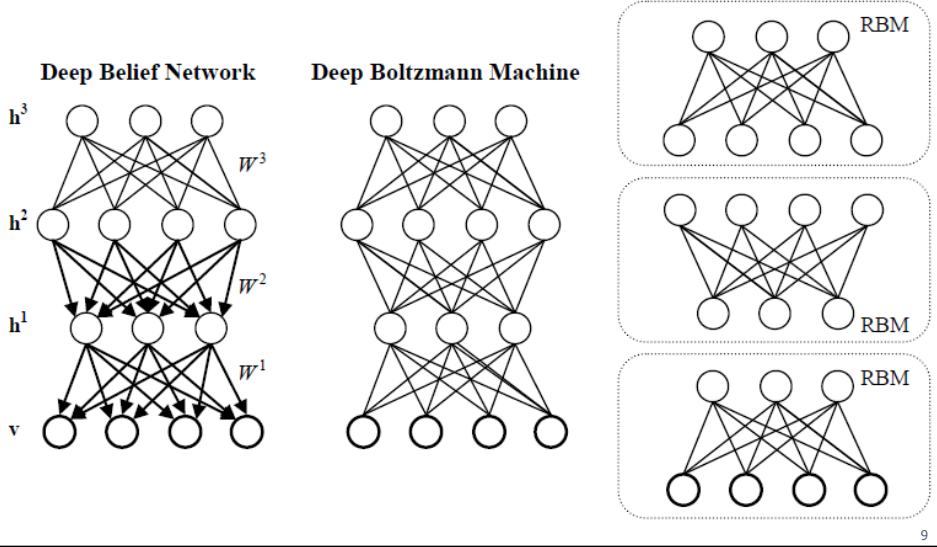
İçerik

- ▶ Derin mimariler
- ▶ Derin inanç ağıları
- ▶ **DBN yapısı**
- ▶ DBN eğitimi
- ▶ DBN parametreleri
- ▶ DBN uygulamaları

8

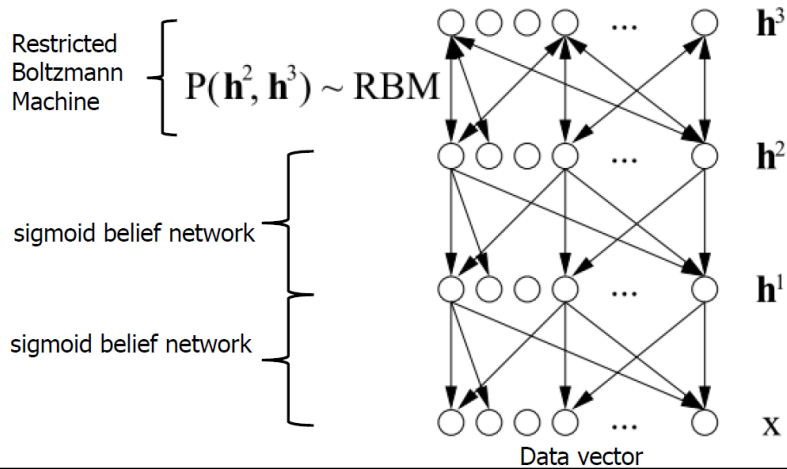
DBN yapısı

- ▶ **Deep belief network, deep Boltzmann machine ve stacked Boltzmann machine.**



DBN yapısı

- ▶ DBN, **bir tane kısıtlanmış Boltzmann makinesi** kullanır.
- ▶ **Diğer katmanlar ileri beslemeli katmanlı** ağ yapısına sahiptir.



İçerik

- ▶ Derin mimariler
- ▶ Derin inanç ağları
- ▶ DBN yapısı
- ▶ **DBN eğitimi**
- ▶ DBN parametreleri
- ▶ DBN uygulamaları

11

DBN eğitimi

- ▶ DBN'lerde **temel problemlerden birisi** parametrelerin **başlangıç değerlerini belirlemektir.**
- ▶ Başlangıç ağırlık değeri uygun olmazsa **gradient descent ile iyi öğrenme yapılamaz.**
- ▶ DBN'lerde, **greedy-layer wise unsupervised learning** kullanılır.
- ▶ Daha iyi sonuç alabilmek için her katmana unsupervised learning yöntemleri ile **pre-train** yapılır.
- ▶ **İlk katmandan başlanarak tüm katmanlar pre-train yapılır.**
- ▶ Tüm katmanlar **pre-train ile initilaze** edilir.
- ▶ Ardından tüm katmanlar **supervised learning** yöntemleri kullanılarak **fine-tuning** yapılır.

12

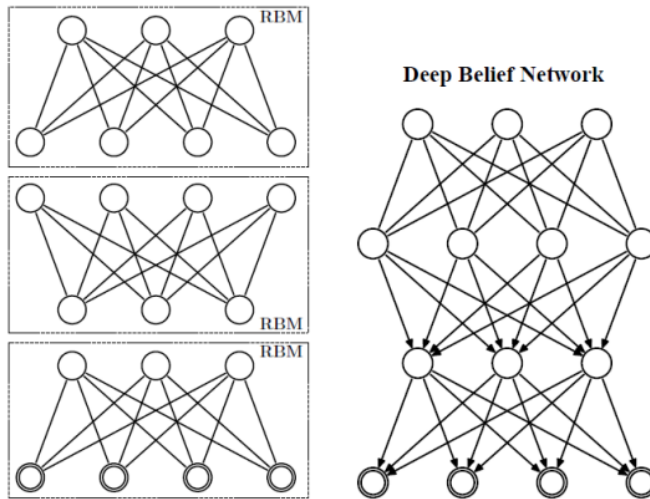
DBN eğitimi

- ▶ DBN'lerde **gizli katmanların her birisi restricted Boltzmann machine (RBM)** olarak alınıp eğitilir.
- ▶ **İlk RBM, input layer ile ilk hidden layer** alınarak oluşturulur.
- ▶ **İlk RBM için eğitim kümesi ilk RBM'in girişidir.**
- ▶ **İkinci RBM, ilk hidden layer ile ikinci hidden layer** alınarak oluşturulur.
- ▶ **İkinci RBM için eğitim kümesi önceki RBM'in çıkışıdır**
- ▶ DBN'lerde **hidden layer sayısı kadar RBM oluşturulur.**
- ▶ Her RBM için ayrı ayrı pre-train yapılır.

13

DBN eğitimi

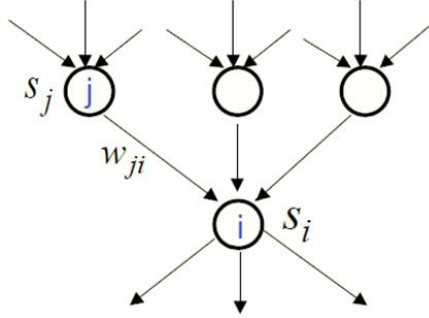
- ▶ Şekilde **solda RBM yığını, sağda** ise ilgili **DBN** görülmektedir.



14

DBN eğitimi

- ▶ **Sigmoid deep belief network** için ağırlıklar **sonraki ve önceki düğümün** durumlarına göre değiştirilir.



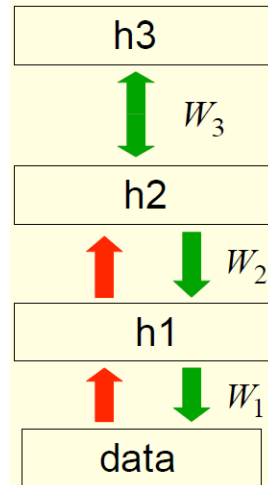
$$p_i = \frac{1}{1 + \exp(-\sum_j s_j w_{ji})}$$

$$\Delta w_{ji} = \eta s_j (s_i - p_i)$$

15

DBN eğitimi

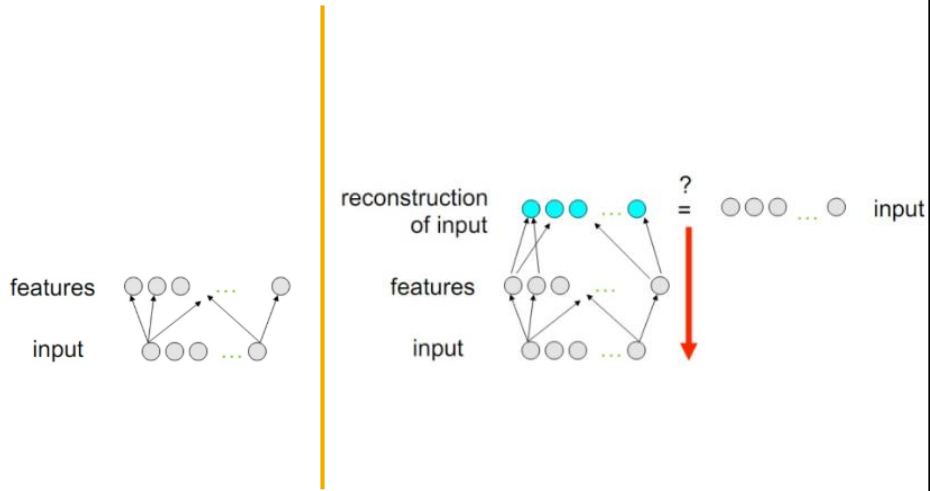
- ▶ Ağırlıkların başlangıç değerleri rastgele atanabilir.
- ▶ DBN'lerde veriyi tekrar üretmek için uzun süreli **Gibbs örnekleme ile en üstteki RBM'ler denge durumuna getirilir.**
- ▶ Ardından **aşağıya doğru geçişle tüm katmanların durumları elde edilir.**
- ▶ Alt katmanlar **veriyi elde etmek için kullanılır.**



16

DBN eğitimi

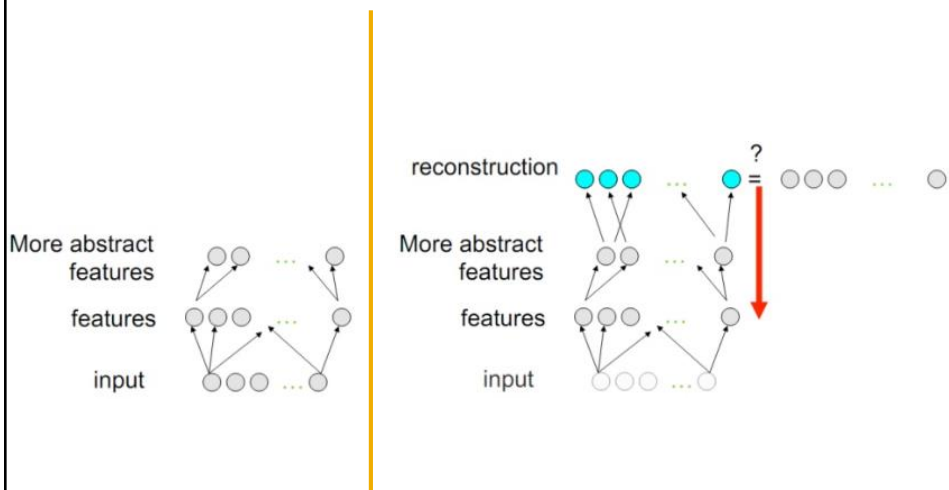
- Her katmanda giriş reconstruction yapılarak parametreler değiştirilir.



17

DBN eğitimi

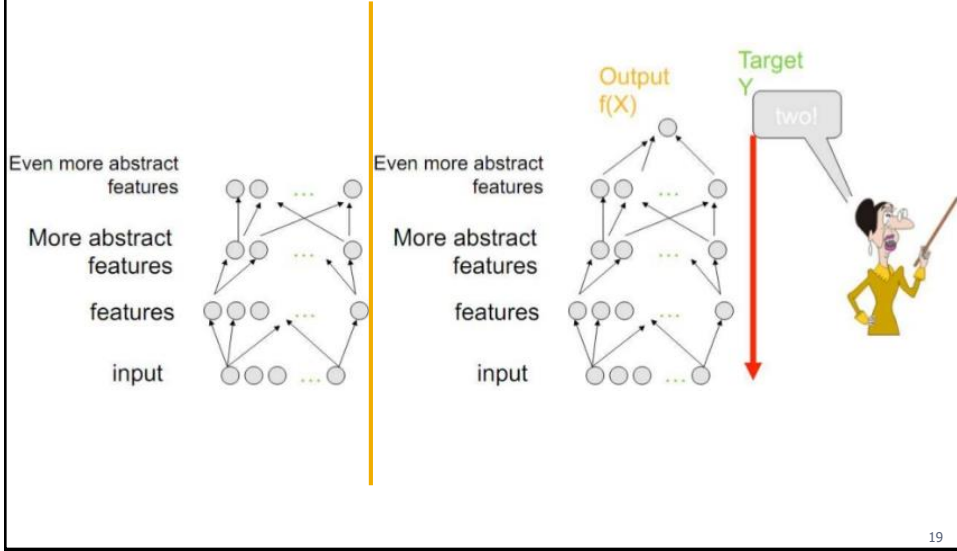
- Her katmanda giriş reconstruction yapılarak parametreler değiştirilir.



18

DBN eğitimi

- Her katmanda giriş reconstruction yapılarak parametreler değiştirilir.



19

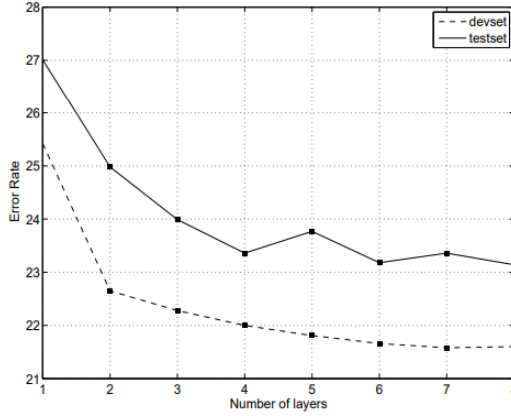
İçerik

- Derin mimariler
- Derin inanç ağları
- DBN yapısı
- DBN eğitimi
- **DBN parametreleri**
- DBN uygulamaları

20

DBN parametreleri

- DBN'lerde **katman sayısı arttıkça hata oranı** başlangıçta hızla **azalmaktadır**.



21

DBN parametreleri

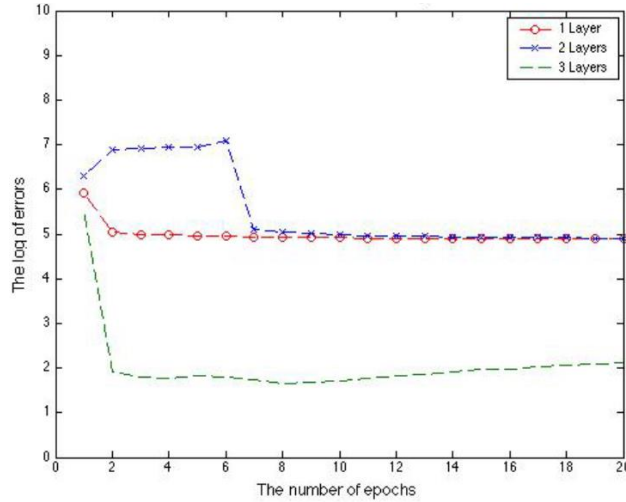
- Her **gizli katmandaki neuron sayısı** hata oranını değiştirmektedir.
- Neuron sayısının artması** hata oranını sürekli **artırmaz** veya **azaltmaz**.
- Tabloda neuron sayısı çok değişse bile hata oranı çok az değişmektedir.
- Tabloda **test kümesinde en düşük hata oranı 2048 neuron** ile elde edilmiştir.

Model	devset	testset
1024 units	21.94%	23.46%
2048 units	22.00%	23.36%
3072 units	21.74%	23.54%

22

DBN parametreleri

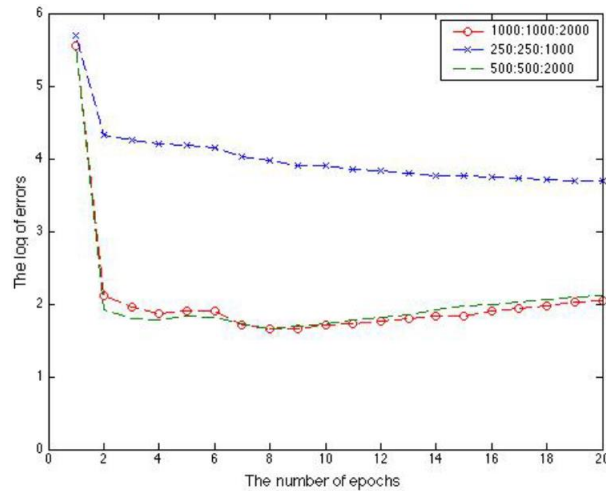
- DBN'de **gizli katman sayısına** bağılı olarak **öğrenme hızı** değişmektedir.



23

DBN parametreleri

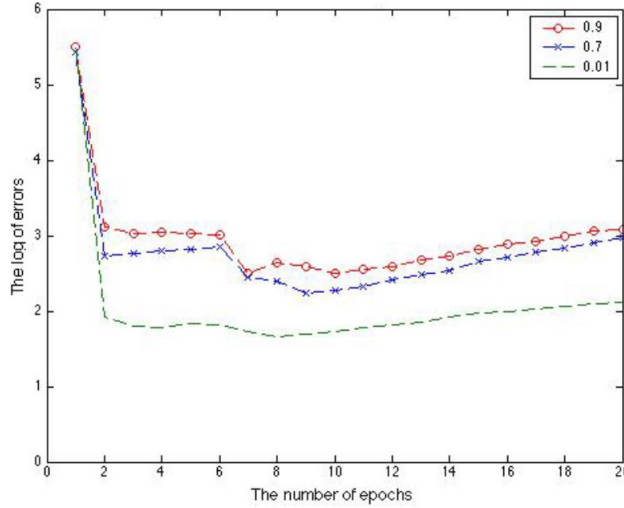
- Gizli katmanlardaki neuron sayısına** bağılı olarak **öğrenme hızı** değişmektedir (Şekilde 3 katmanlı DBN için değişim verilmiştir.).



24

DBN parametreleri

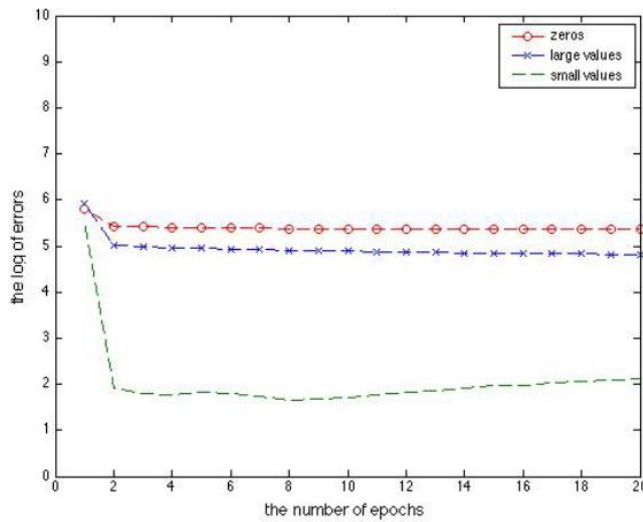
- Öğrenme katsayısına (η) bağlı olarak **öğrenme hızı** değişmektedir.



25

DBN parametreleri

- Ağırlıkların **başlangıç değerleri** öğrenme hızını **değiştirmektedir**.



26

İçerik

- ▶ Derin mimariler
- ▶ Derin inanç ağıları
- ▶ DBN yapısı
- ▶ DBN eğitimi
- ▶ DBN parametreleri
- ▶ **DBN uygulamaları**

27

DBN uygulamaları

- ▶ DBN günümüzde aşağıdaki başlıca alanlarda uygulanmaktadır:
 - ▶ **Kümeleme**
 - ▶ **Sınıflandırma**
 - ▶ **Boyut indirgeme**
 - ▶ **Anormallik tespiti**
 - ▶ **Doğal dil işleme**

28