

Derin Öğrenme Deep Learning

Hazırlayan: M. Ali Akcayol
Gazi Üniversitesi
Bilgisayar Mühendisliği Bölümü

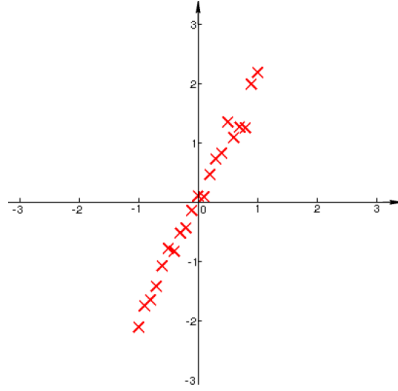
Bu dersin sunumları, "Simon Haykin, Neural Networks and Learning Machines, Prentice Hall, 2016." kitabı kullanılarak hazırlanmıştır.

İçerik

- ▶ Autoencoder
- ▶ AE'lerin yapısı
- ▶ AE'larda ileri geçiş
- ▶ AE'lerin eğitimi
- ▶ AE hiper parametreleri
- ▶ AE uygulamaları

Autoencoder

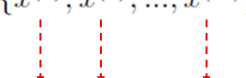
- ▶ Cep telefonundaki veriyi **minimum ağ kaynağı** kullanarak bulut ortamına göndermek isteyelim.
- ▶ **Minimum bant genişliği kullanmak için veriyi optimize etmemiz gerekir.**
- ▶ Cep telefonundaki veri aşağıdaki gibi iki boyutlu olsun.



3

Autoencoder

- ▶ Veri noktalarındaki **dikey eksen değeri yatay eksen değerinin yaklaşık iki katıdır.**
- ▶ Ağ üzerinden **sadece yatay eksen değerini gönderip** bulutta **2 ile çarparak dikey eksen değerini hesaplayabiliriz.**
- ▶ Bu durumda sıkıştırmadan dolayı belirli bir **kayıp oluşur**, ancak **ağ trafiği %50 azalır.**
- ▶ **Çok boyutlu verilerde** sıkıştırma yapmak için **daha az boyutlu bir kümeye eşleştirme** gereklidir.
- ▶ **n boyutlu m tane veri olan x kümesini, k boyutlu m tane veri olan z kümesine dönüştürecek yöntem** geliştirilir ($n > k$).

$$\{x^{(1)}, x^{(2)}, \dots, x^{(m)}\}$$

$$\{z^{(1)}, z^{(2)}, \dots, z^{(m)}\}$$

4

Autoencoder

- ▶ Cep telefonundaki **her x verisi z kümesinde bir elemana eşleştirilir.**
- ▶ z kümesindeki eleman buluta gönderilir.
- ▶ Bulutta z kümesindeki eleman x kümesindeki bir elemana eşleştirilir.

$$\begin{array}{c} \{z^{(1)}, z^{(2)}, \dots, z^{(m)}\} \\ \downarrow \quad \downarrow \quad \downarrow \\ \{x^{(1)}, x^{(2)}, \dots, x^{(m)}\} \end{array}$$

- ▶ **Sistematik eşleştirme** için aşağıdaki eşitlikler kullanılabilir.

$$z^{(i)} = W_1 x^{(i)} + b_1$$

$$\tilde{x}^{(i)} = W_2 z^{(i)} + b_2$$

- ▶ Burada, W_1 , W_2 , b_1 , ve b_2 değerleri **belirlenmelidir.**

5

Autoencoder

- ▶ Burada **amaç**, $\tilde{x}^{(i)}$ değerini $x^{(i)}$ değerine yaklaştırmaktır.
- ▶ Aşağıdaki **maliyet fonksiyonu minimize edilmelidir.**

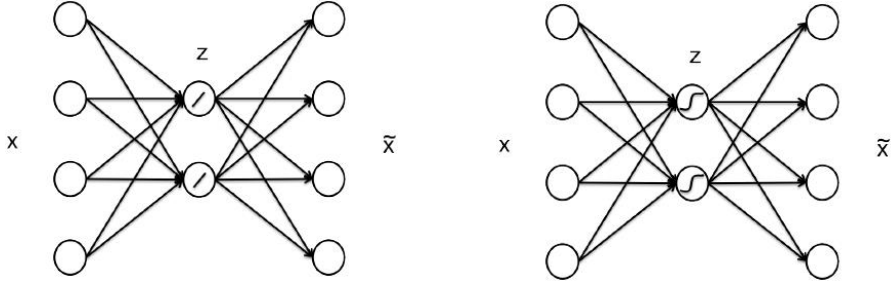
$$\begin{aligned} J(W_1, b_1, W_2, b_2) &= \sum_{i=1}^m \left(\tilde{x}^{(i)} - x^{(i)} \right)^2 \\ &= \sum_{i=1}^m \left(W_2 z^{(i)} + b_2 - x^{(i)} \right)^2 \\ &= \sum_{i=1}^m \left(W_2 (W_1 x^{(i)} + b_1) + b_2 - x^{(i)} \right)^2 \end{aligned}$$

- ▶ Amaç fonksiyonunu minimize eden W_1 , W_2 , b_1 ve b_2 değerleri **stochastic gradient descent** ile bulunabilir.

6

Autoencoder

- ▶ Oluşturulan **bu mimari autoencoder olarak adlandırılır.**
- ▶ Kullanılan **aktivasyon fonksiyonuna göre lineer** veya **nonlineer autoencoder** oluşturulabilir.



7

İçerik

- ▶ Autoencoder
- ▶ AE'lerin yapısı
- ▶ AE'larda ileri geçiş
- ▶ AE'lerin eğitimi
- ▶ AE hiper parametreleri
- ▶ AE uygulamaları

8

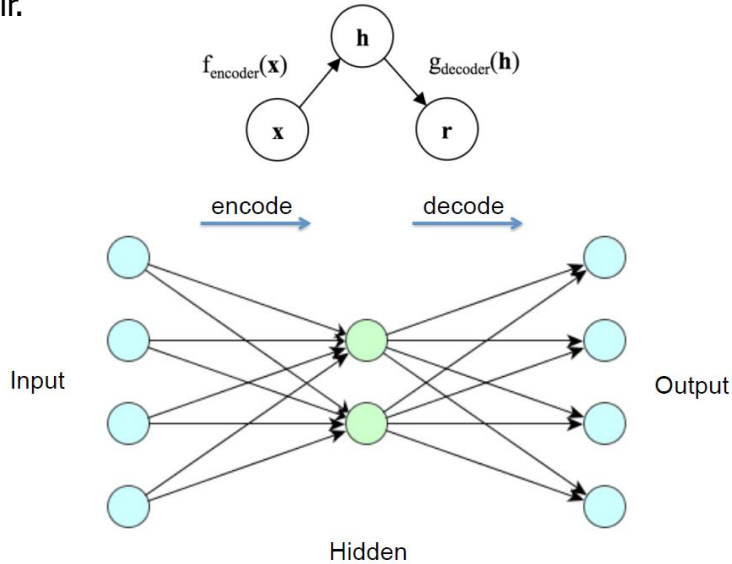
AE'ların yapısı

- ▶ **Autoencoder (AE)**, aldığı **girişi** kendi **çıkışına kopyalamak amacıyla** eğitilen yapay sinir ağıdır.
- ▶ Autoencoder, **multi layered feedforward ANN'lerin özel bir türüdür.**
- ▶ AE'da girişi temsil eden bir gizli katman vardır (h).
- ▶ Basit bir AE, **encoder** ve **decoder** olarak **iki parçadan oluşur:**
 - ▶ **Encoder fonksiyonu** $h = f(x)$
 - ▶ **Decoder fonksiyonu** $r = g(h)$
- ▶ **Öğrenme sürecinde** $r = x$ yapılmaya çalışılır.
- ▶ x girişlerine karşılık y etiketleri yoktur, **autoencoder unsupervised learning yapar.**

9

AE'ların yapısı

- ▶ Eğer bir **AE başarılı tasarlanmış ise**, $g(f(x)) = x$ kümesini öğrenir.



10

AE'ların yapısı

- ▶ AE'daki gizli katman **h 'nin boyutu (code size) x 'in boyutundan küçükse (undercomplete)**, girişteki **belirgin özellikleri öğrenir**.
- ▶ AE'daki gizli katman **h 'nin boyutu (code size) x 'in boyutuna eşit veya büyükse (overcomplete)**, belirgin özelliklerin yanında diğer özellikleri de öğrenir.
- ▶ Öğrenme sürecinde, **kayıp fonksiyonu minimize edilir**.

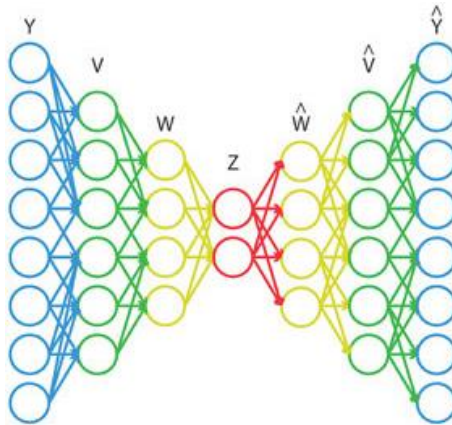
$$L(x, g(f(x)))$$

- ▶ Kayıp fonksiyonu **L , hata ölçümü yapan bir fonksiyondur** (Örn., hataların karelerinin ortalaması).

11

AE'ların yapısı

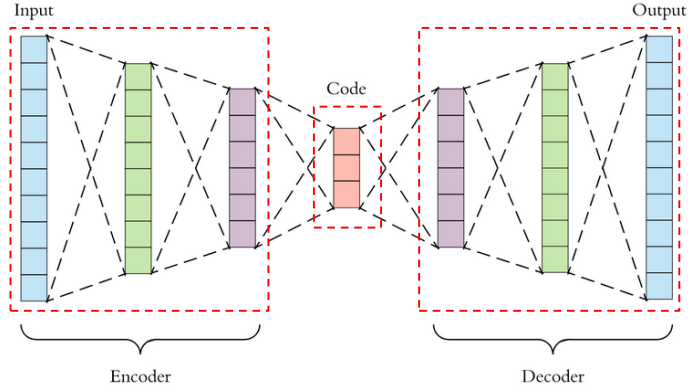
- ▶ Çok sayıda gizli katman kullanılarak derin autoencoder geliştirilebilir.



12

AE'lerin yapısı

- ▶ AE, **encoder** ile katmanlardaki **neuron sayılarını azaltır**, daha sonra **decoder** ile katmanlardaki **neuron sayılarını artırır**.



13

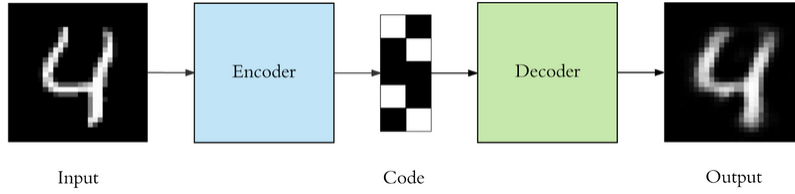
İçerik

- ▶ Autoencoder
- ▶ AE'lerin yapısı
- ▶ AE'larda ileri geçiş
- ▶ AE'lerin eğitimi
- ▶ AE hiper parametreleri
- ▶ AE uygulamaları

14

AE'larda ileri geiř

- ▶ **Gizli katmanda** elde edilen **code**, **giriřin zeti** veya **sıkıřtırılmıř halidir.**
- ▶ **Encoder**, **giriři sıkıřtırır** ve **code retir.**
- ▶ **Decoder** ise, aldıđı **code** ile **giriři yeniden elde etmeye alıřır.**
- ▶ Bir AE oluřturmak iin, **encoding metodu**, **decoding metodu** ve **kayıp fonksiyonu** gereklidir.



15

AE'larda ileri geiř

- ▶ AE, giriř verisinde **boyut indirgemek amacıyla kullanılabilir.**
- ▶ Giriř verilerini **gizli katmanda daha kk boyut ile ifade eder.**
- ▶ **ıkıř verisi, giriř verisinden** her zaman **farklıdır.**
- ▶ ıkıř etiketleri iin ek bilgiye ihtiya duymaz, **unsupervised ğrenme yapar.**
- ▶ **ıkıř etiketlerini** eđitim verisinden (giriřlerden) **kendisi oluřturduđu iin self-supervised** olarak adlandırılır.

16

İçerik

- ▶ Autoencoder
- ▶ AE'lerin yapısı
- ▶ AE'larda ileri geçiş
- ▶ **AE'lerin eğitimi**
- ▶ AE hiper parametreleri
- ▶ AE uygulamaları

17

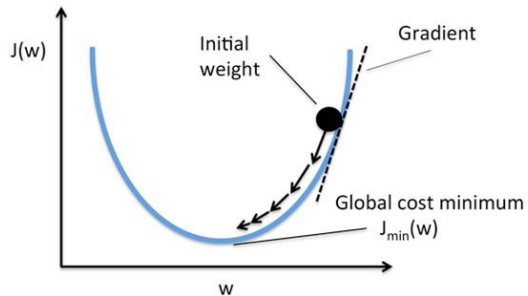
AE'lerin eğitimi

- ▶ AE'lerin eğitimi için **ağırlıklar gradient descent ile değiştirilir.**
- ▶ AE, eğitim yaparken veriye göre kendi etiketlerini kendisi üretir.
- ▶ AE'larda maliyet fonksiyonunu minimize eden **ağırlıklar hesaplanır.**

$$J(\mathbf{w}) = \frac{1}{2} \sum_i (\text{target}^{(i)} - \text{output}^{(i)})^2$$

$$\Delta w_j = -\eta \frac{\partial J}{\partial w_j}$$

$$\mathbf{w} := \mathbf{w} + \Delta \mathbf{w}_j$$



18

AE'lerin eğitimi

- ▶ Çok büyük miktardaki verilerin eğitiminde **stochastic gradient descent** kullanılır.
- ▶ **Gradient descent** ile her örnek veri girişi için ağırlıklar değiştirilir.

- for one or more epochs, or until approx. cost minimum is reached:
 - for training sample i :
 - for each weight j
 - $w_j := w + \Delta w_j$, where: $\Delta w_j = \eta(\text{target}^{(i)} - \text{output}^{(i)})x_j^{(i)}$

- ▶ **Stochastic gradient descent** ile birden fazla epoch için hesaplanan hata değerine göre **ağırlıklar değiştirilir**.

- for one or more epochs:
 - for each weight j
 - $w_j := w + \Delta w_j$, where: $\Delta w_j = \eta \sum_i (\text{target}^{(i)} - \text{output}^{(i)})x_j^{(i)}$

19

İçerik

- ▶ Autoencoder
- ▶ AE'lerin yapısı
- ▶ AE'larda ileri geçiş
- ▶ AE'lerin eğitimi
- ▶ **AE hiper parametreleri**
- ▶ AE uygulamaları

20

AE hiper parametreleri

- ▶ **Kod boyutu**
 - ▶ Kod boyutu küçüldükçe veri sıkıştırma oranı artar.
 - ▶ Veri sıkıştırma oranı arttıkça kayıp miktarı artar.
- ▶ **Katman sayısı**
 - ▶ Katman sayısı veri miktarına ve karmaşıklığına göre seçilir.
- ▶ **Katmanlardaki düğüm sayısı**
 - ▶ Kodlayıcı kısmındaki katmanlardaki düğüm sayıları her katmanda azalır.
 - ▶ Kod çözücü kısmında ise katmanlardaki düğüm sayıları her katmanda artar.
- ▶ **Kayıp fonksiyonu**
 - ▶ Kayıp fonksiyonu için genellikle mean squared error (MSE) kullanılır.
 - ▶ Ancak, probleme özgü farklı kayıp fonksiyonları belirlenebilir.

21

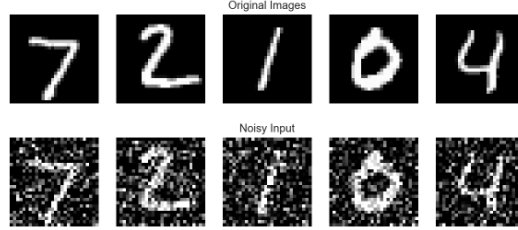
İçerik

- ▶ Autoencoder
- ▶ AE'lerin yapısı
- ▶ AE'larda ileri geçiş
- ▶ AE'lerin eğitimi
- ▶ AE hiper parametreleri
- ▶ **AE uygulamaları**

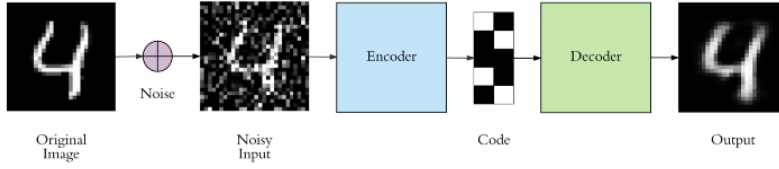
22

AE uygulamaları

- ▶ Denoising AE, gürültülü veride gürültüyü yok etmek için kullanılır.



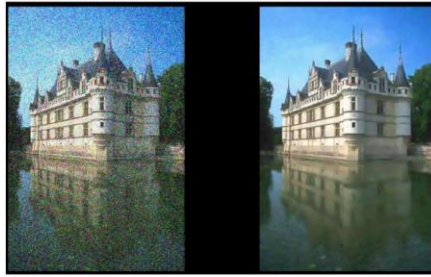
- ▶ Elde edilen sonuç orijinalden farklıdır, ancak gürültü yok edilir.



23

AE uygulamaları

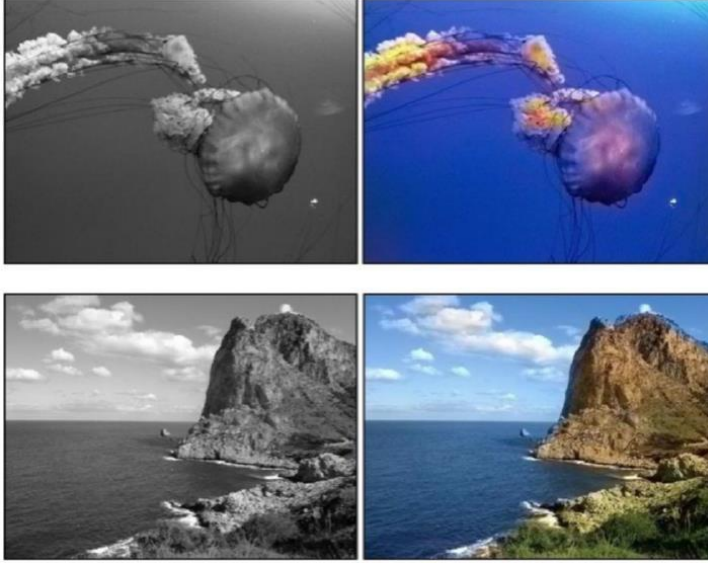
- ▶ Veriyi gürültüden temizlemek amacıyla kullanılır.



24

AE uygulamaları

- ▶ Görüntüyü renklendirme amacıyla kullanılmaktadır.



25

AE uygulamaları

- ▶ Varyasyonel AE, convolution+deconvolution katmanlarıyla birlikte kullanılarak sentetik insan yüzü oluşturabilmektedir*.

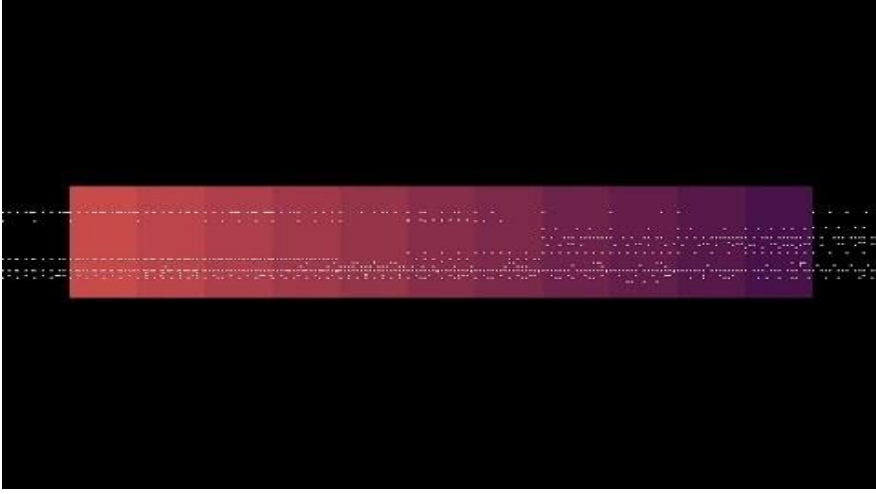


* <https://github.com/yzwxx/vae-celebA>

26

AE uygulamaları

- Varyasyonel AE ile sentetik müzik oluşturabilmektedir*.



* <https://www.youtube.com/watch?v=OsxP4IFp76I>

27

Ödev

- AE'lerin veri artırmaya (data augmentation) yönelik uygulamasını içeren bir makale hakkında ödev hazırlayınız.

28