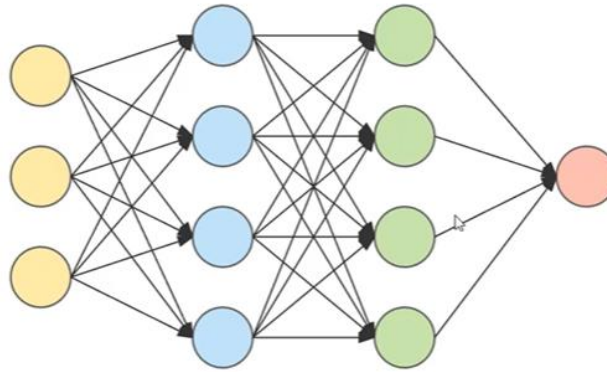


$$y = W \times x + b$$

$$y = W \times x + b$$

Giriş Katmanı



1. Gizli Katman

2. Gizli Katman

Çıkış Katmanı

EĞİTİCİLİ (GÖZETİMLİ) ÖĞRENME

SINIFLANDIRMA (CLASSIFICATION) YÖNTEMLERİ

- Destek Vektör Makineleri (Support Vector Machines)
- Ayırıştırma Analizi (Discriminant Analysis)
- Naif Bayes (Naive Bayes)
- En Yakın Komşu (Nearest Neighbor)

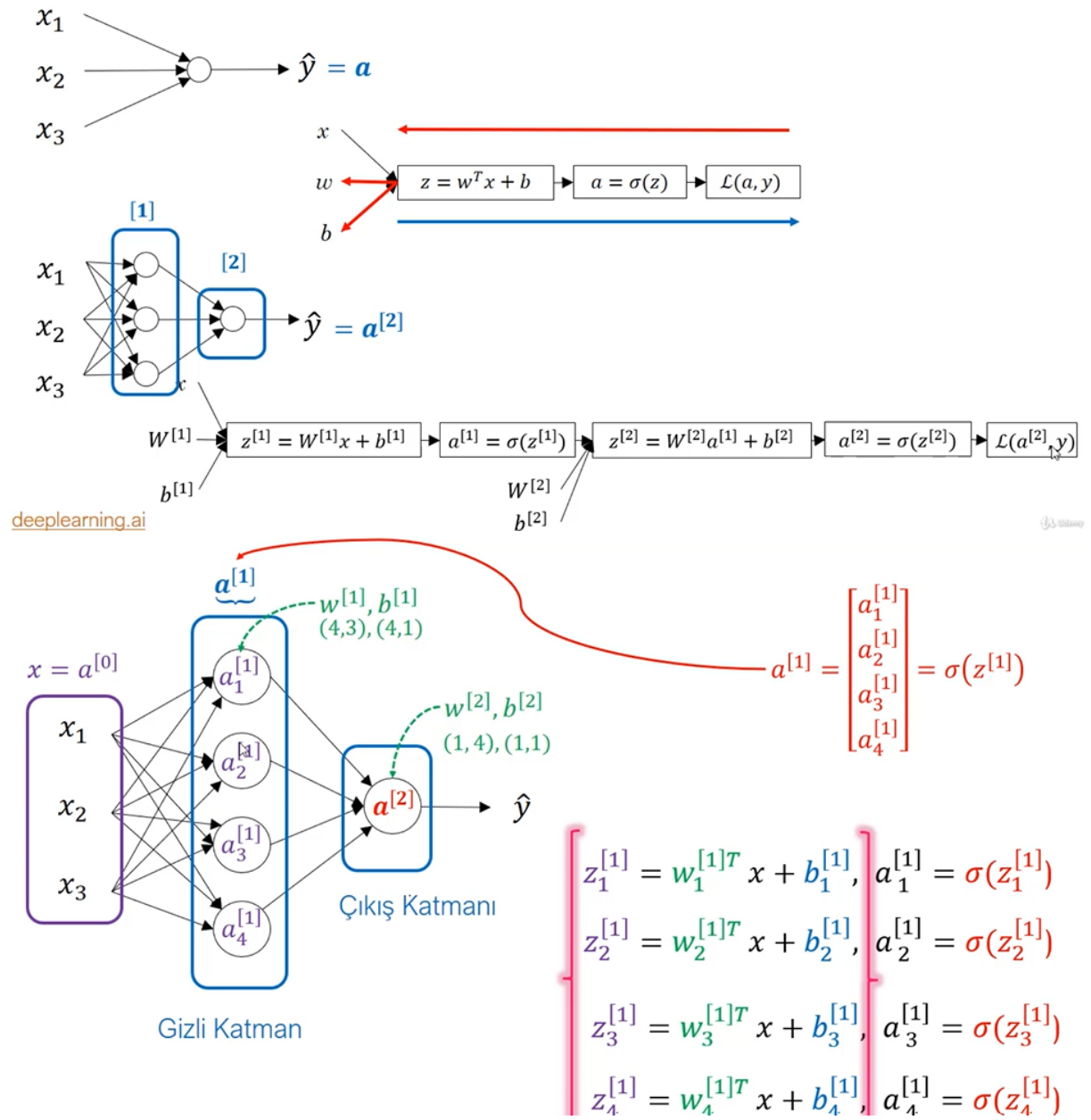
BAĞLANIM (REGRESSION) YÖNTEMLERİ

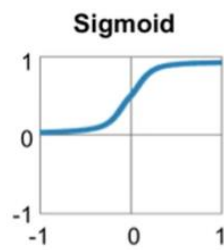
- Lineer Bağlanım
- Karar Ağaçları (Decision Trees)
- Yapay Sinir Ağları

EĞİTİCİSİZ (GÖZETİMSİZ) ÖĞRENME

KÜMELEME YÖNTEMLERİ

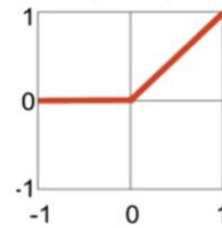
- K-Ortalama, Bulanık C-Ortalama
- Gauss Karışımı (Gaussian Mixture)
- Saklı Markov Modeli (Hidden Markov Model)
- Yapay Sinir Ağları





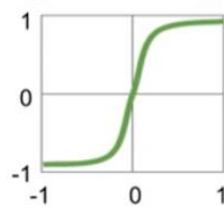
$$\text{Sigmoid: } a = \frac{1}{1 + e^{-z}}$$

Rectified Linear Unit (ReLU)



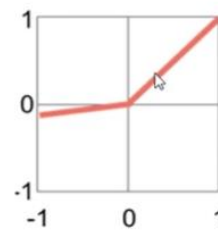
$$\text{ReLU: } a = \max(0, z)$$

Hyperbolic Tangent



$$\text{Tanh: } a = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

Leaky ReLU



$$\text{Leaky ReLU: } a = \max(0.01 \cdot z, z)$$

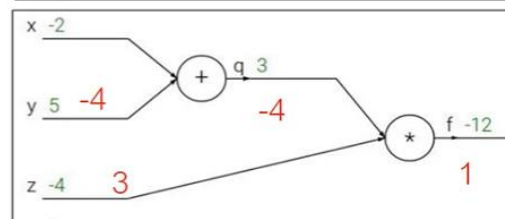
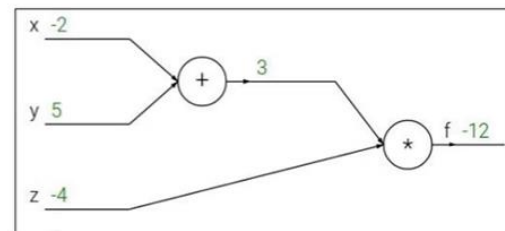
$$f(x, y, z) = (x + y)z$$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1 \quad \frac{\partial q}{\partial y} = 1$$

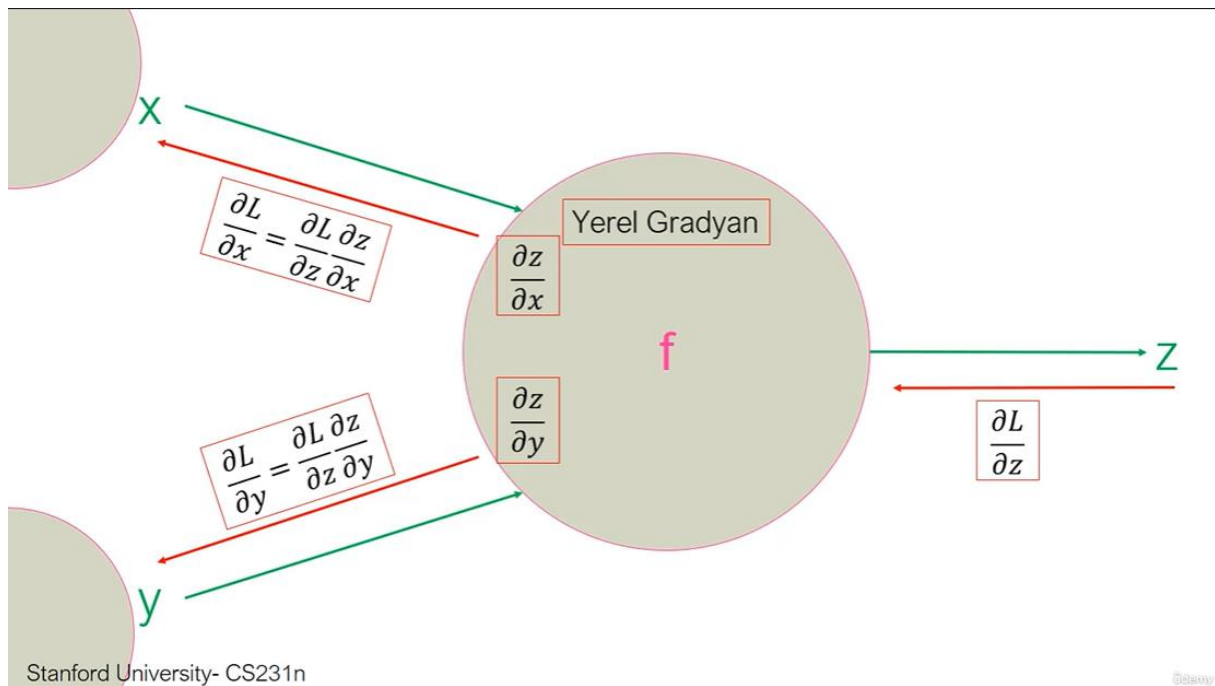
$$f = qz \quad \frac{\partial f}{\partial q} = z \quad \frac{\partial f}{\partial z} = q$$

Zincir Kuralı

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial q} \frac{\partial q}{\partial x}$$

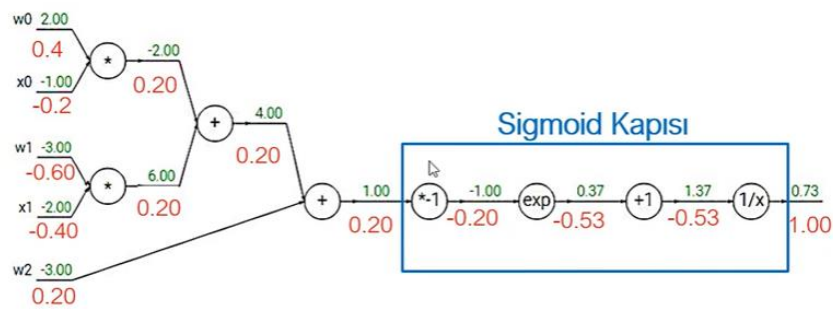


$$\frac{\partial f}{\partial x}$$



$$f(x) = \frac{1}{1 + e^{-(w_0 x_0 + w_1 x_1 + w_2 x_2)}}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$



$$\frac{d\sigma(x)}{dx} = \frac{e^{-x}}{(1 + e^{-x})^2} = \left(\frac{1 + e^{-x} - 1}{1 + e^{-x}} \right) \left(\frac{1}{1 + e^{-x}} \right) = (1 - \sigma(x))\sigma(x)$$

$$= (1 - 0.73) \cdot 0.73 = 0.20$$

Parametreler: $w^{[1]}, b^{[1]}, w^{[2]}, b^{[2]}$

$$n_x = n^{[0]}, n^{[1]}, n^{[2]} = 1$$

$$(n^{[1]}, n^{[0]}) (n^{[1]}, 1) (n^{[2]}, n^{[1]}) (n^{[2]}, 1)$$

Maliyet Fonksiyonu: $J(w^{[1]}, b^{[1]}, w^{[2]}, b^{[2]}) = \frac{1}{m} \sum_{i=1}^n L(\hat{y} - y)$

Gradyan iniş algoritması: $dw^{[1]} = \frac{dJ}{dw^{[1]}}, db^{[1]} = \frac{dJ}{db^{[1]}} \dots$

$$w^{[1]} = w^{[1]} - \alpha \cdot dw^{[1]}$$

$$b^{[1]} = b^{[1]} - \alpha \cdot db^{[1]}$$

...

İleri Yayılım Algoritması
(Feed Forward Propagation)

$$Z^{[1]} = w^{[1]}x + b^{[1]}$$

$$A^{[1]} = g^{[1]}(Z^{[1]})$$

$$Z^{[2]} = w^{[2]}A^{[1]} + b^{[2]}$$

$$A^{[2]} = g^{[2]}(Z^{[2]}) = \sigma(Z^{[2]})$$

Geriye Yayılım Algoritması
(Back Propagation)

$$dZ^{[2]} = A^{[2]} - Y$$

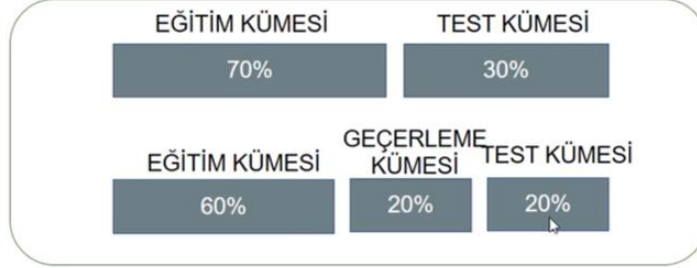
$$dW^{[2]} = \frac{1}{m} dZ^{[2]} A^{[1]T}$$

$$dZ^{[1]} = \underbrace{W^{[2]T} dZ^{[2]}}_{(n^{[1]}, m)} * \underbrace{g^{[1]'}(Z^{[1]})}_{(n^{[1]}, m)}$$

$$dW^{[1]} = \frac{1}{m} dZ^{[1]} X^T$$

Test/Geçerleme/Eğitim Kümesi Seçimi

Gelecekte elde etmeyi beklediğiniz verileri yansıtacak bir geçerleme kümesi ve test kümesi seçin.



Karışıklık Matrisi/
Olasılık Tablosu

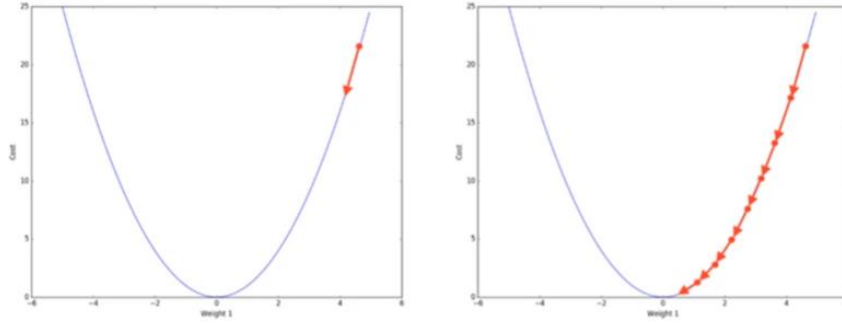
Kestirim Sınıfı (Beklenen)	Gerçek Sınıflar (Gözlem)	
	tp (true positive) Gerçek Pozitif	fp (false positive) Beklenmedik Sonuç
	fn (false negative) Kaybedilen Sonuç	tn (true negative) Gerçek Negatif

$$\text{Precision (Hasasiyet)} = \frac{tp}{tp+fp}$$

$$\text{Accuracy (Doğruluk)} = \frac{tp+tn}{tp+tn+fp+fn}$$

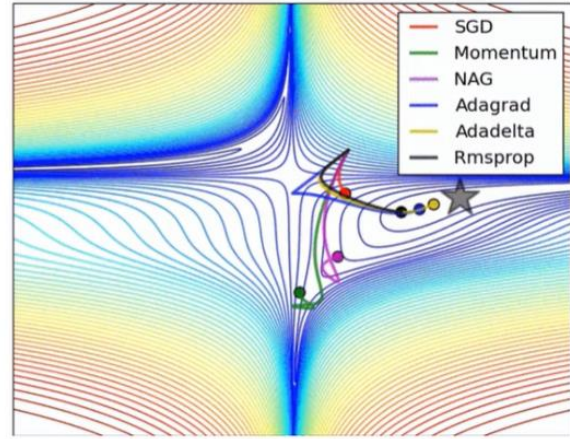
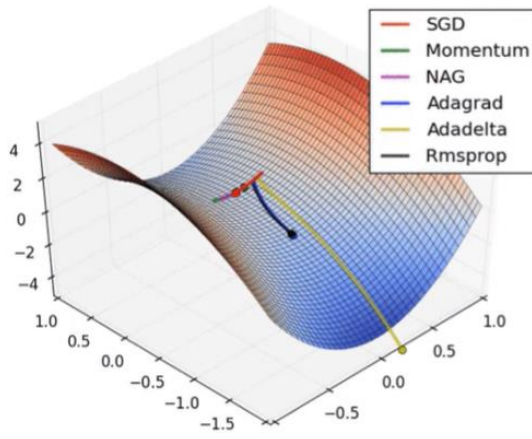
$$\text{Recall (Kesinlik)} = \frac{tp}{tp+fn}$$

ÖĞRENME: OPTİMİZASYON PROBLEMİ

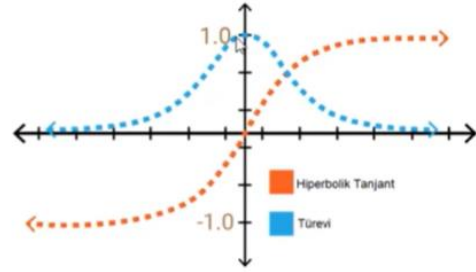
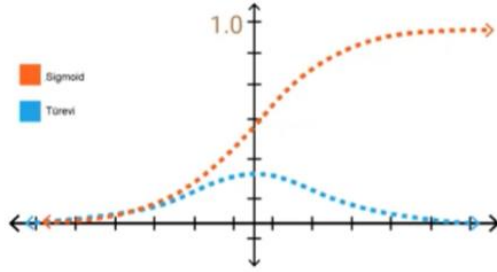


Hedef: Ağırlıkları ve biası güncelleyerek yitim fonksiyon çıkışı minimize etmek

Önerilen teknik: Küçük-Kümeleme gradyan iniş ya da stokastik gradyan iniş algoritması kullanılabilir.



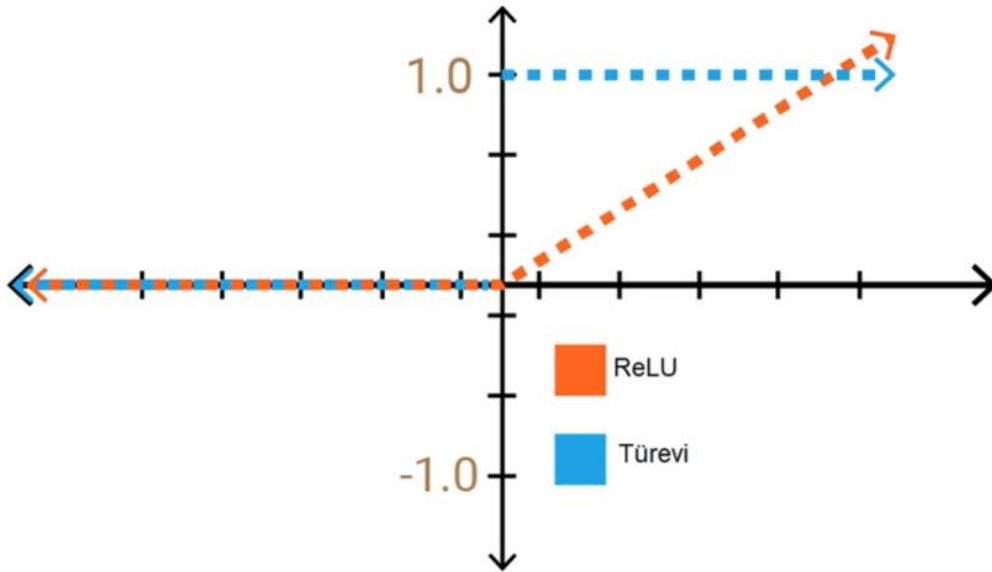
AKTİVASYON FONKSİYONU SEÇİMİ

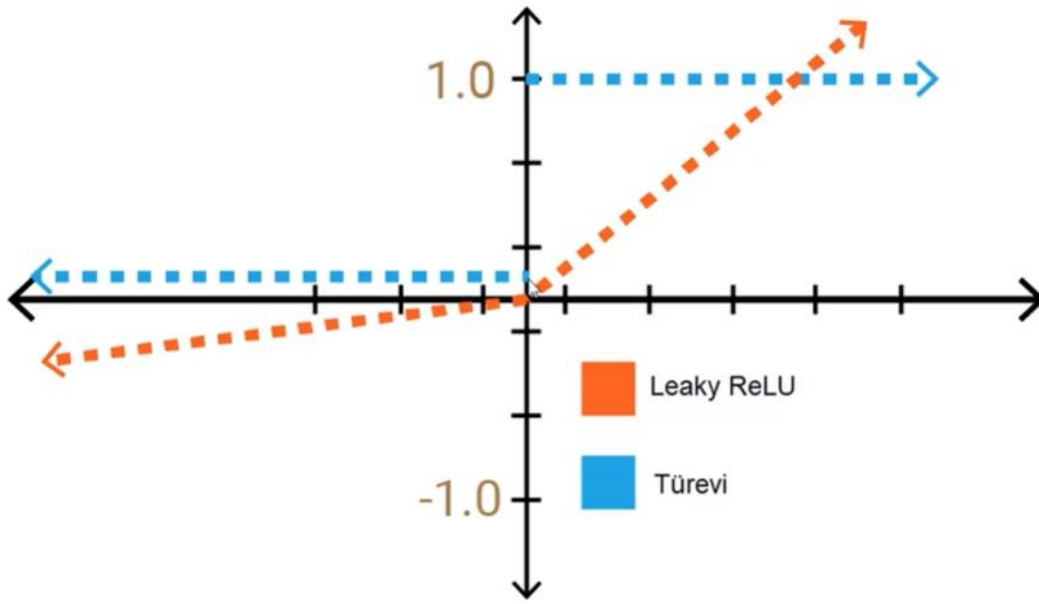


$$\frac{d\sigma(x)}{dx} = (1 - \sigma(x)) \sigma(x)$$



Parçalı türev adımları küçük seçildiğinde = Öğrenme yavaşlar

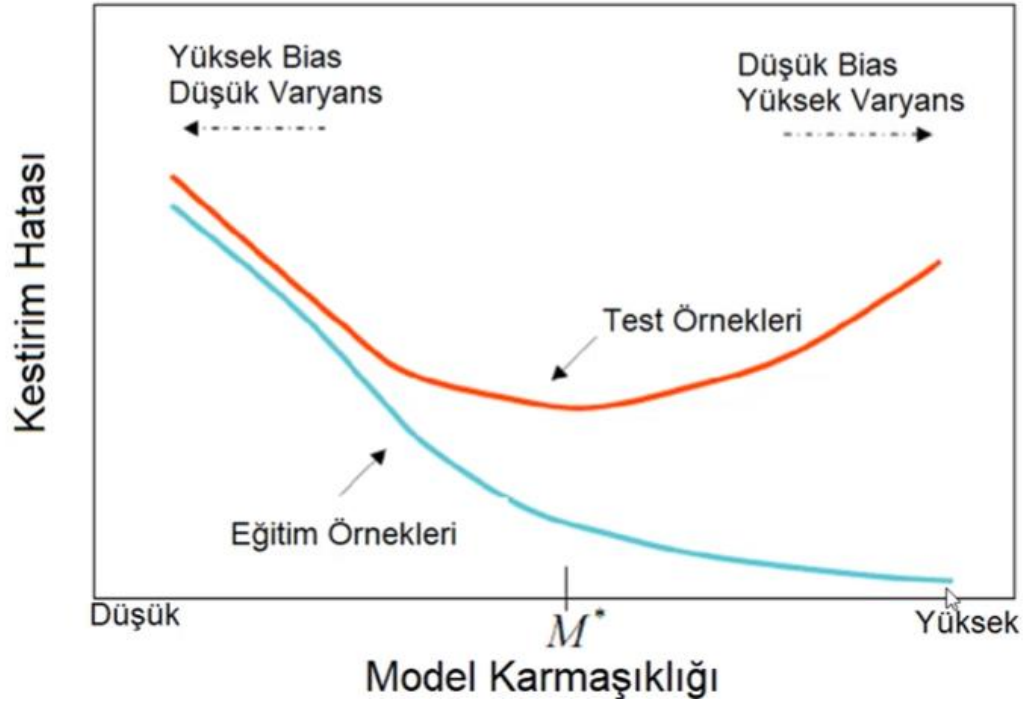




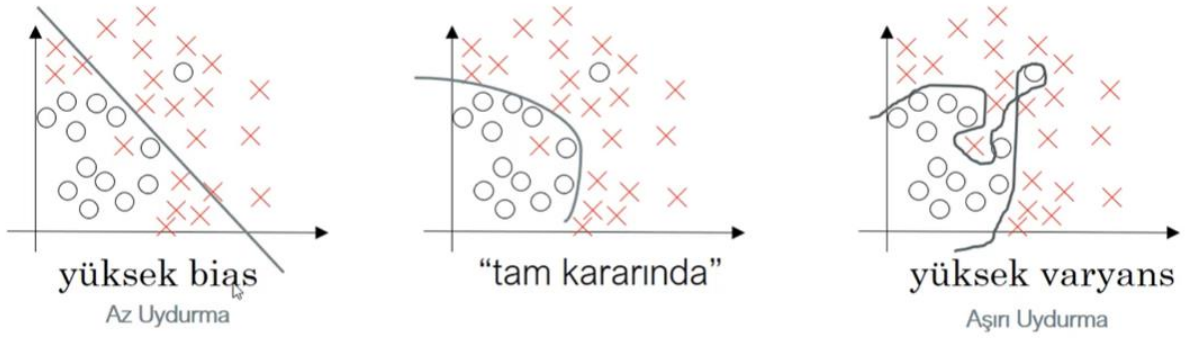
EĞİTİM TUR SAYISI (Epoch)

- Epoch sayısı arttıkça başarımlar bir süre artmaya devam eder.
- Başarımların artması azaldığında ya da durduğunda daha fazla epoch işlem yapmaya gerek yoktur.
- Bu gibi hiper parametreler eğitim süresini optimize eder.
- Epoch sayısı problemin türüne göre değişkendir.

Bias vs. Variance



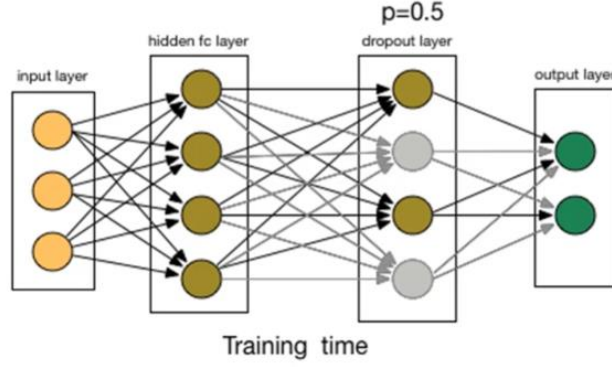
Bias vs. Variance



Bias = modelin tahminlerindeki sistematik hata.

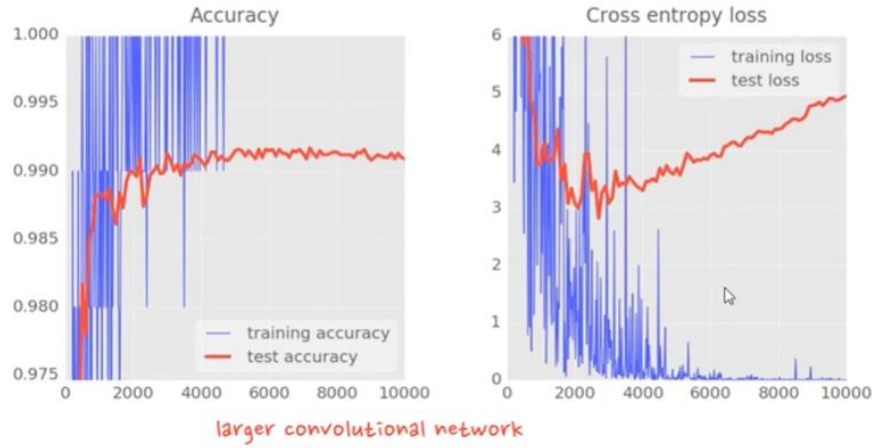
Variance = eğitim setinde örnekleme gürültüsü tahminlerdeki gürültüye ne neden olur.

Düğüm Seyreltme (Dropout)



Hedef: Modeli genelleştirmeye yardımcı olur.

Düğüm seyreltme uygulandığı ve uygulanmadığı durumlarda başarımların değişim örneği



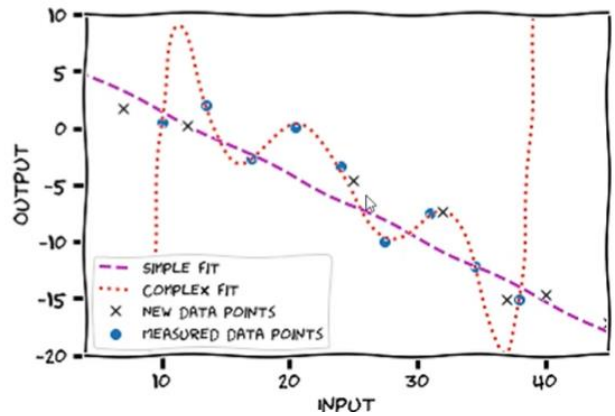
Gürültü Ekleme İşlemi

Modelin girdilerine gürültü ekleme

Modelin gizli katmanlarına gürültü ekleme

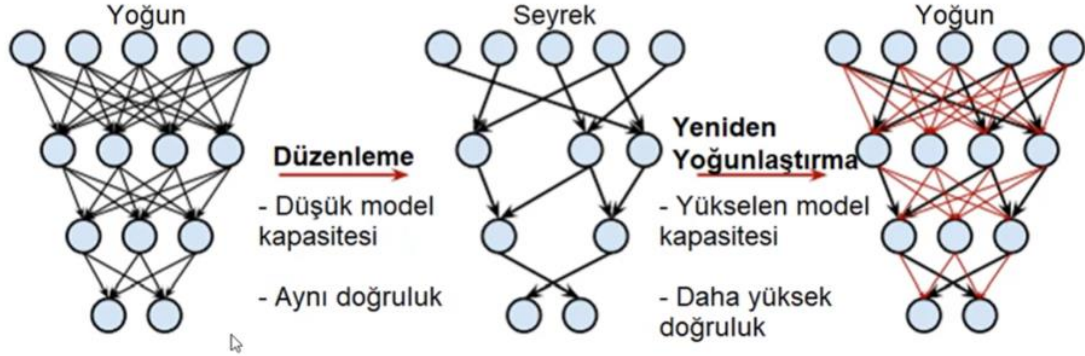
Modelin ağırlık parametrelerine
gürültü ekleme:
Recurrent Neural Network (RNN)

Yitim fonksiyonuna, gürültüyü parametre
olarak ekleme



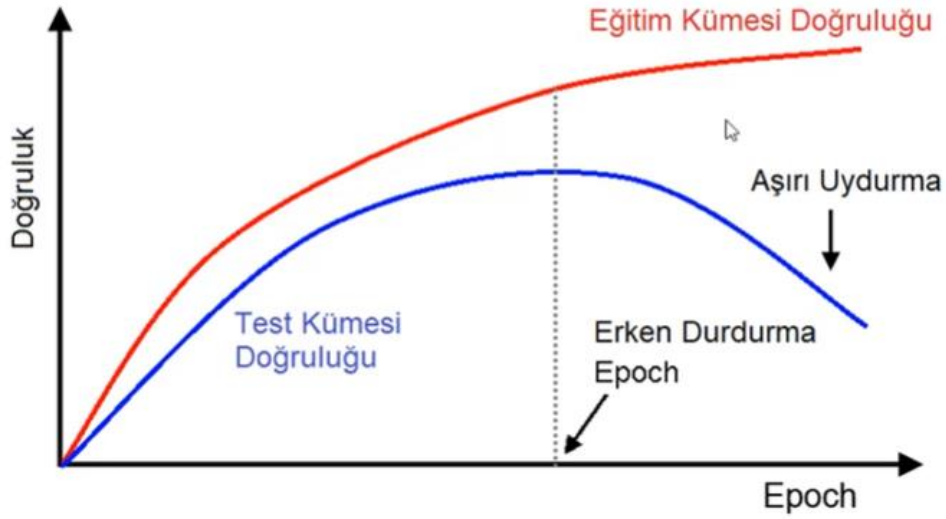
Yoğun-Seyrek-Yoğun Ağ ile Eğitim

(Dense-Sparse-Dense Training)

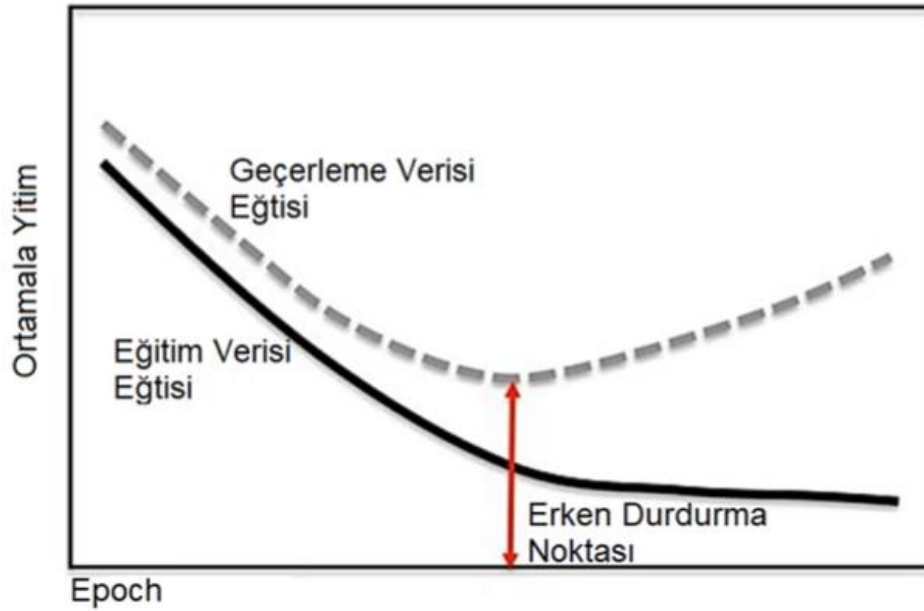


Aşırı uydurma- Az uydurma

(Overfitting vs. Underfitting)



Aşırı Uydurmadan Önce Durdurma (Early Stopping)



TRANSFER ÖĞRENME (TRANSFER LEARNING)

