

Karmaşık Sistemler ve Veri Bilimi
Yaz Okulu 2019

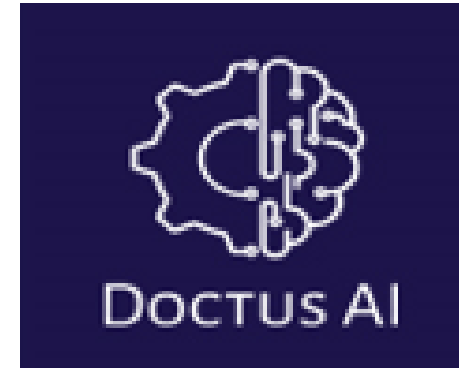
DERİN ÖĞRENME

M. AYYÜCE KIZRAK



@ayyucekizrak

www.ayyucekizrak.com





BLACK MIRROR
BANDERSNATCH

MAVİ HAP

KIRMIZI HAP

Karmaşık Sistemler ve Veri Bilimi Yaz Okulu 2019-M. Ayyüce
Kızrak



DERİN ÖĞRENME VE VERİ İLİŞKİSİ

- Derin öğrenme nedir?
- Yapay sinir ağları çalışma yapısı
- Veriye olan ihtiyacın boyutu



TEMEL DERİN ÖĞRENME MODELLERİ

- Evrişimli sinir ağları
- Özyinelemeli sinir ağları
- Üretici çekişmeli ağlar



DERİN ÖĞRENME VE KARMAŞIK SİSTEMLER

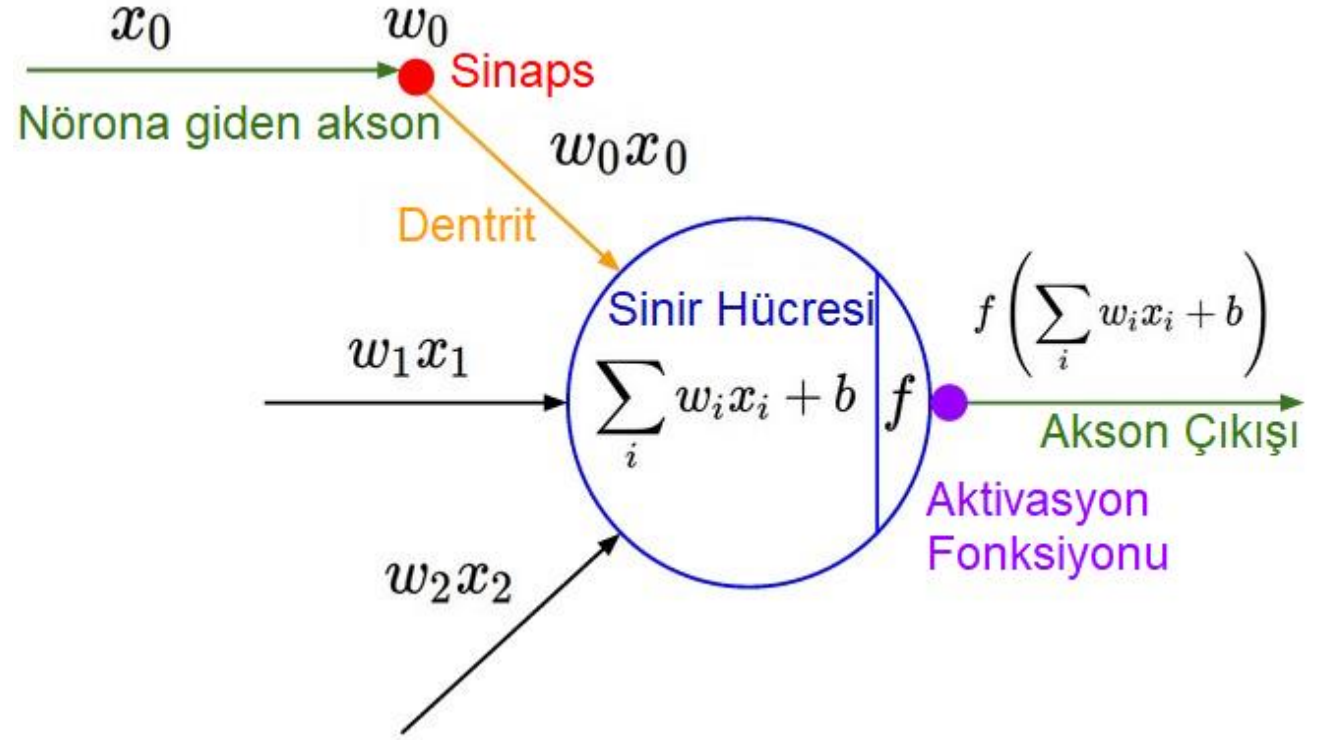
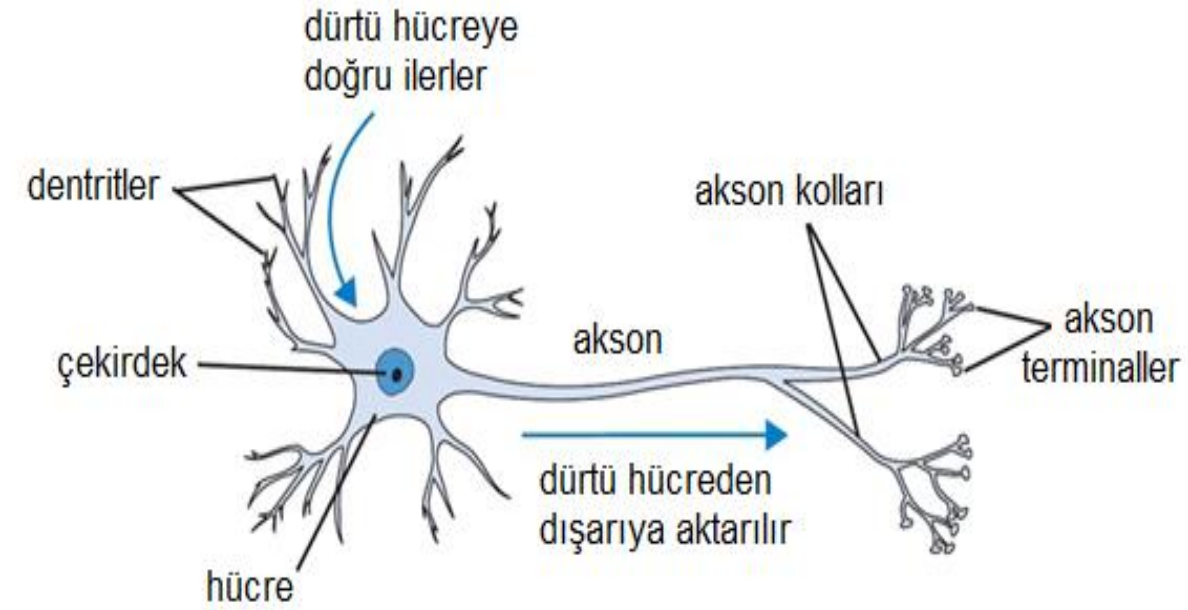
- Katmanlar ve katmanlarda elde edilen özellikler
- Doğrusal olmayan fonksiyonların katmanlara etkisi
- Neden doğrusal ve lojistik regresyon yöntemlerinden daha başarılı



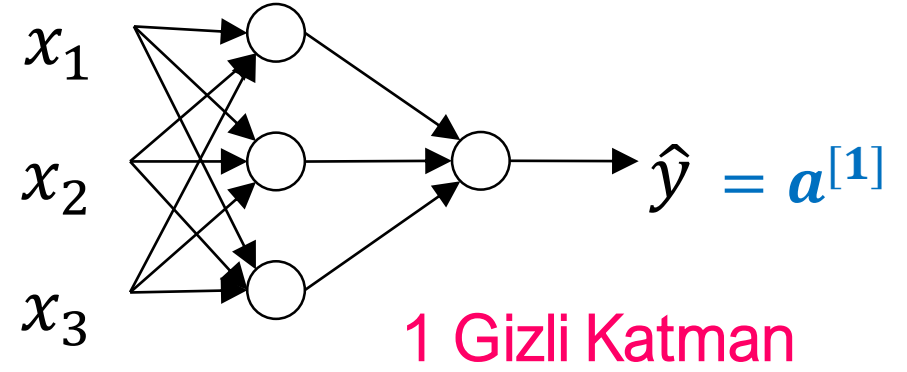
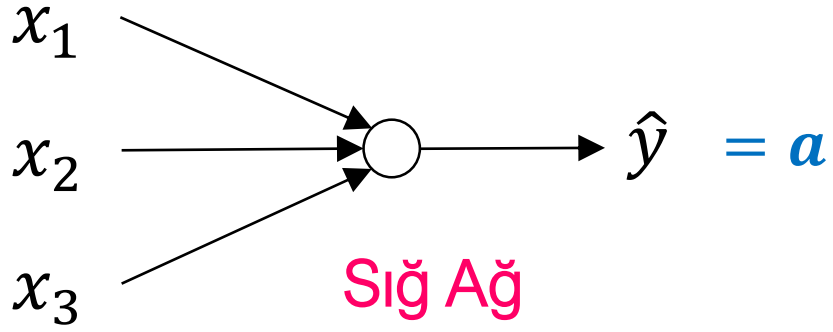
DERİN ÖĞRENMEDE ESKİ-YENİ YAKLAŞIMLAR



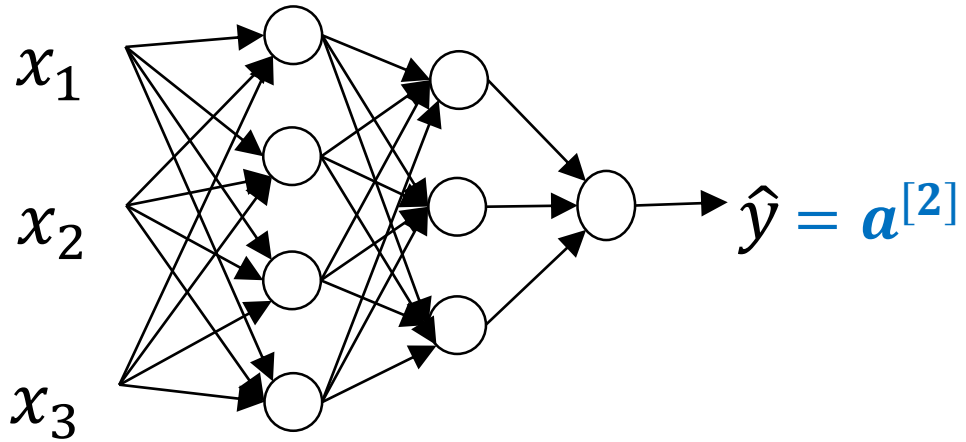
DERİN ÖĞRENME & VERİ İLİŞKİSİ



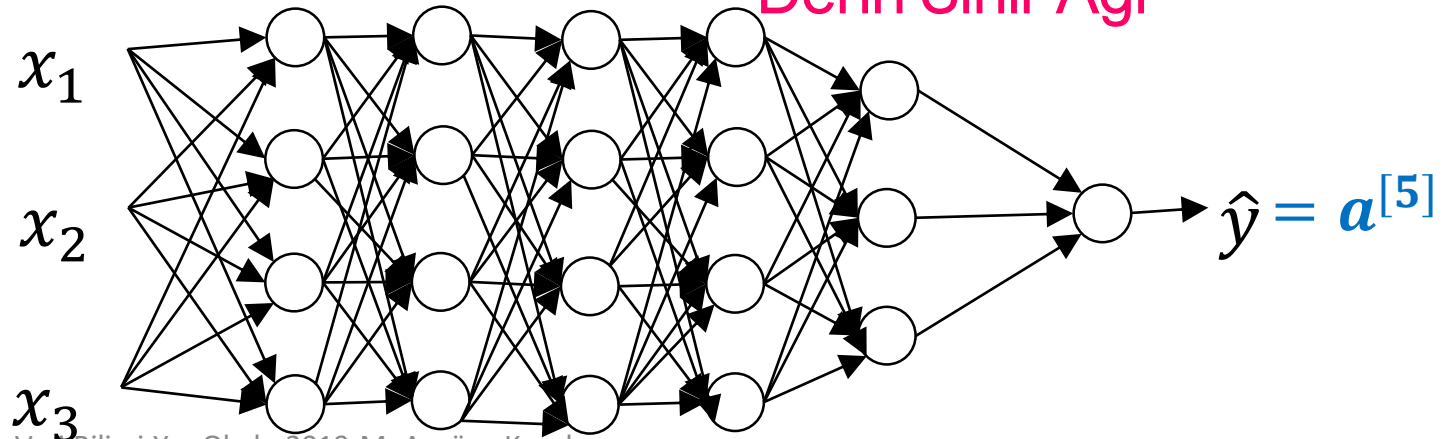
$$y = W \times x + b$$

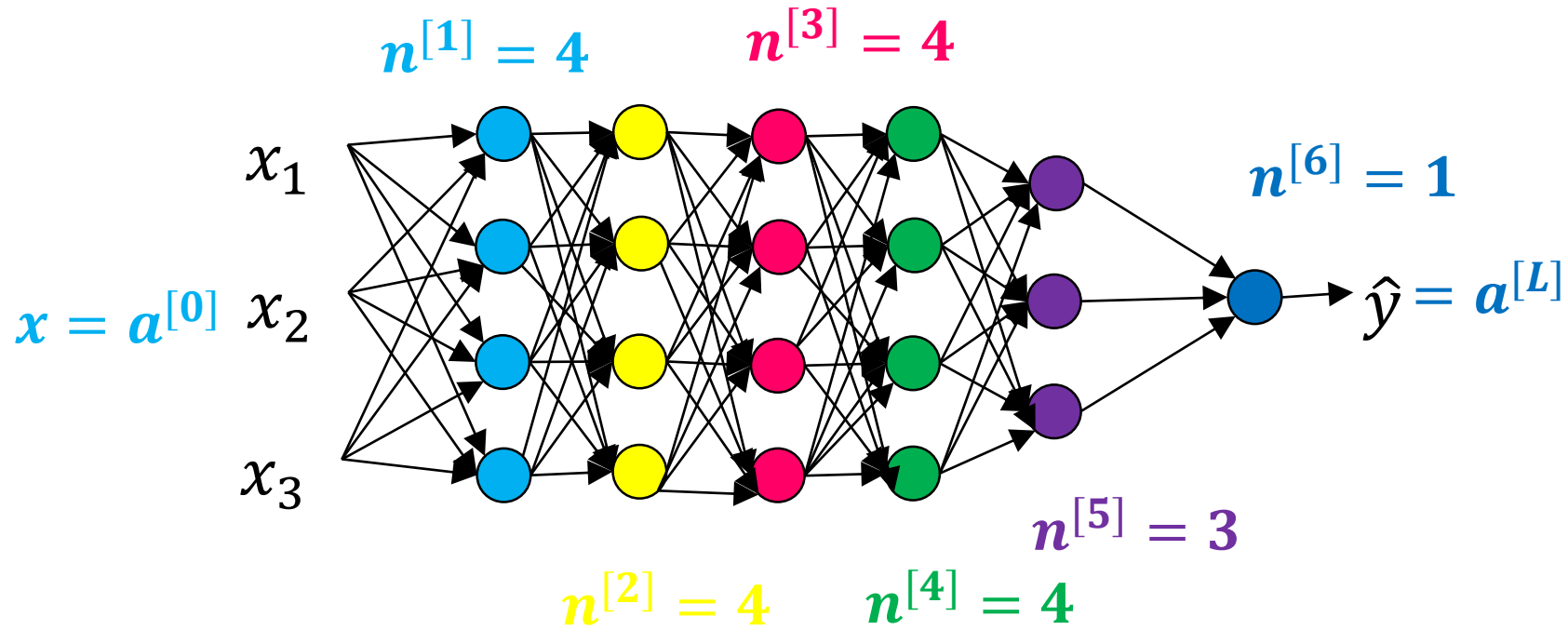


2 Gizli Katman



5 Gizli Katmanlı
Derin Sinir Ağı





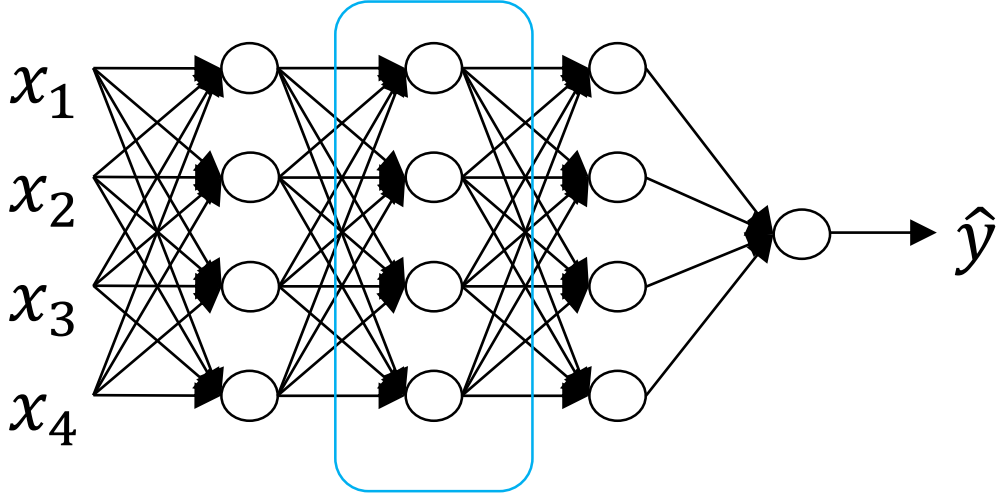
$L = \text{Katman Sayısı} = 6$

$n^{[l]} = l. \text{ katmandaki nöron sayısı}$

$a^{[l]} = g^{[l]}(z^{[l]})$, $w^{[l]} = l. \text{ katmandaki ağırlıklar}$

Karmaşık Sistemler ve Veri Bilimi Yazı Okulu 2019-M. Ayvüce Kızrak

L. Katman için



İleri yönde:

Giriş: $a^{[l-1]}$, Çıkış: $a^{[l]}$, Ara: $z^{[l]}$

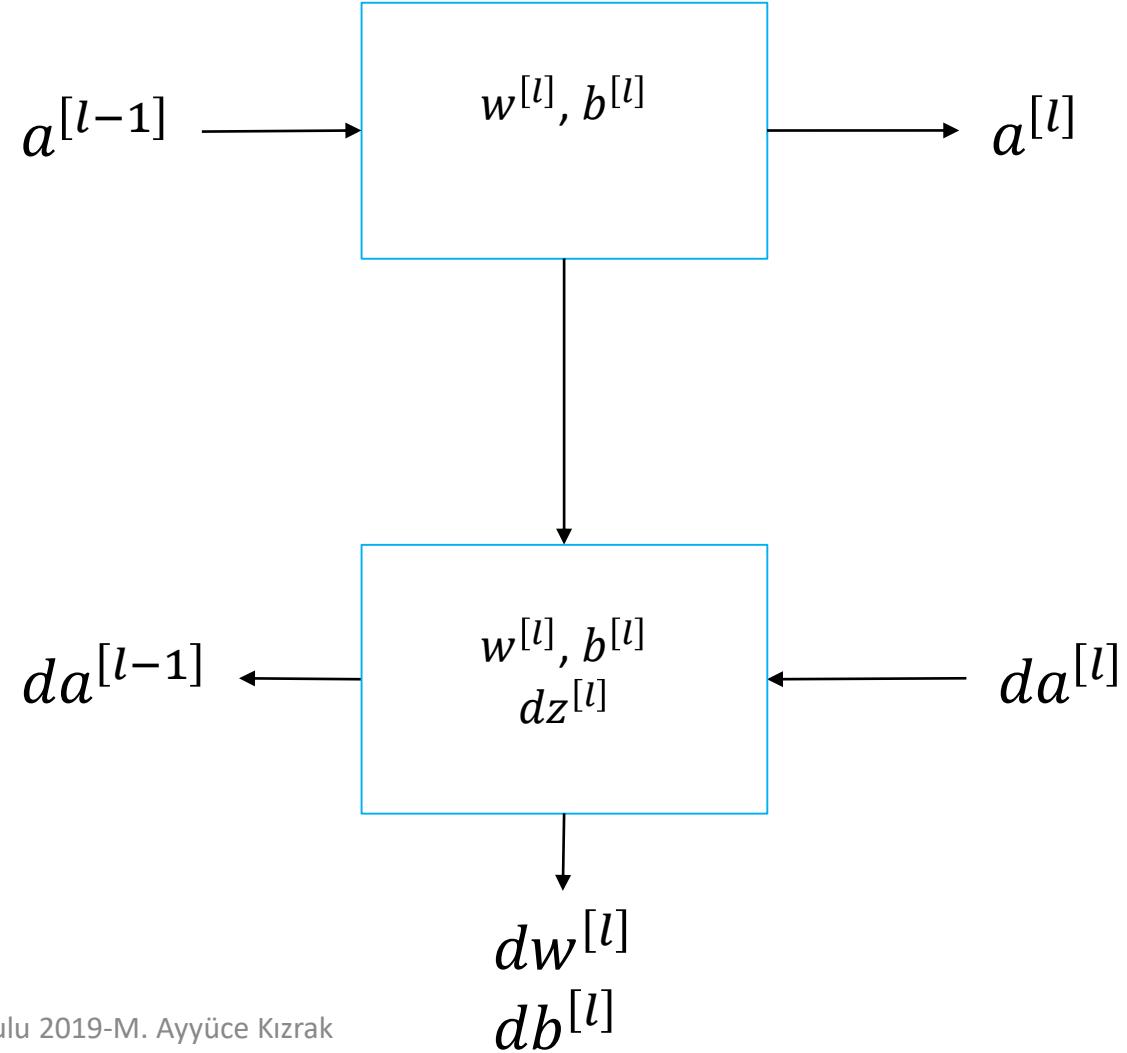
$$z^{[l]} = w^{[l]} \cdot a^{[l-1]} + b^{[l]}$$

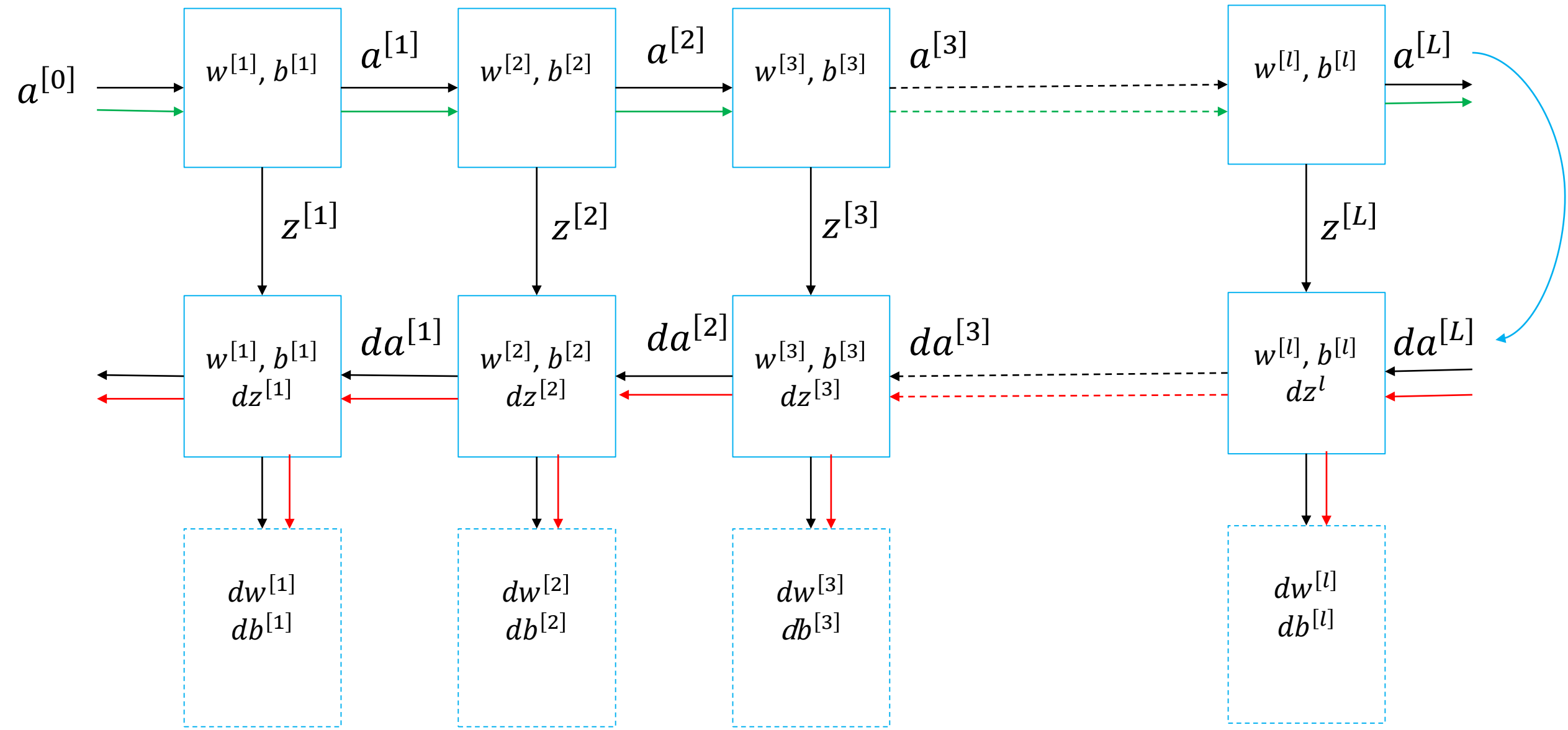
$$a^{[l]} = g^{[l]}(z^{[l]})$$

Geri yönde:

Giriş: $da^{[l]}$, Çıkış: $da^{[l-1]}$, $dw^{[l]}$, $db^{[l]}$ Ara: $z^{[l]}$

L. Katman için





Karmaşık Sistemler ve Veri Bilimi Yaz Okulu 2019-İl. Ayşe Kızrak

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

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

L. Katman İleri Yayılımda:

Giriş: $a^{[l-1]}$

Çıkış: $a^{[l]}$, 0 Ara: $z^{[l]}$

$$z^{[l]} = w^{[l]} \cdot a^{[l-1]} + b^{[l]}$$

$$a^{[l]} = g^{[l]}(z^{[l]})$$

L. Katman Geri Yayılımda:

Giriş: $da^{[l]}$

Çıkış: $da^{[l-1]}$, $dw^{[l]}$, $db^{[l]}$

$$dz^{[l]} = da^{[l]} * g^{[l]}(z^{[l]})$$

$$dw^{[l]} = dz^{[l]} * a^{[l-1]}$$

$$db^{[l]} = dz^{[l]}$$

$$da^{[l-1]} = w^{[l]T} * dz^{[l]}$$

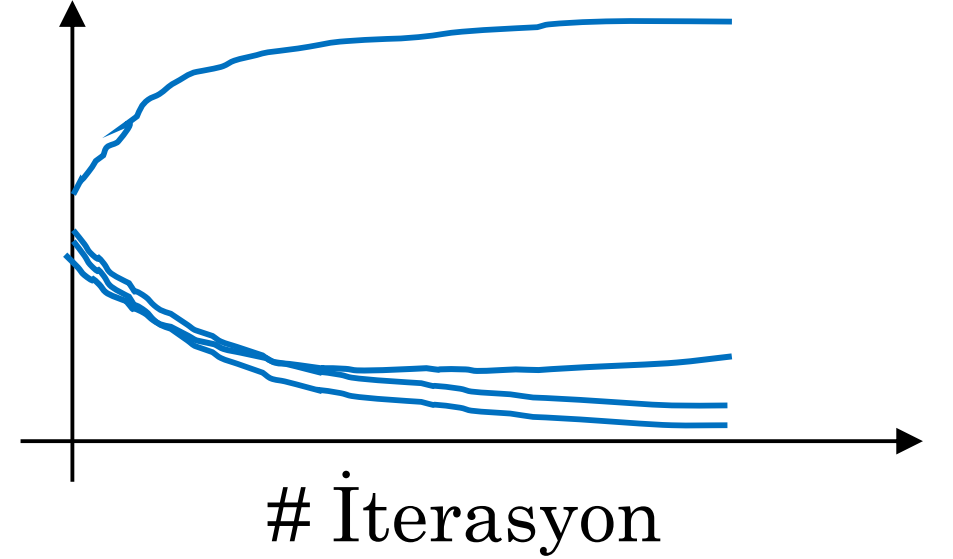
$$dz^{[l]} = w^{[l+1]T} dz^{[l+1]} * g^{[l]}(z^{[l]})$$

PARAMETRELER: $W^{[1]}, b^{[1]}, W^{[2]}, b^{[2]}$...

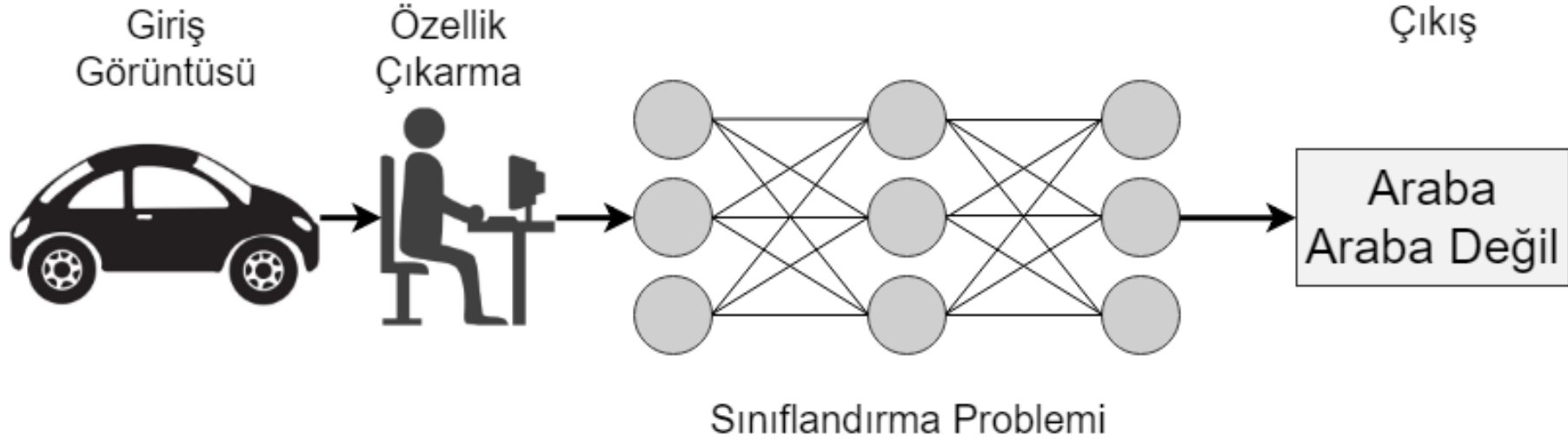
HİPER PARAMETRELER:

- *Öğrenme Oranı (Learning Rate)*
- *İterasyon Sayısı*
- *Gizli Katman Sayısı*
- *Aktivasyon Fonksiyonu Seçimi*
- *Küme Boyutu ve diğer düzenleme (regularization) yöntemleri*

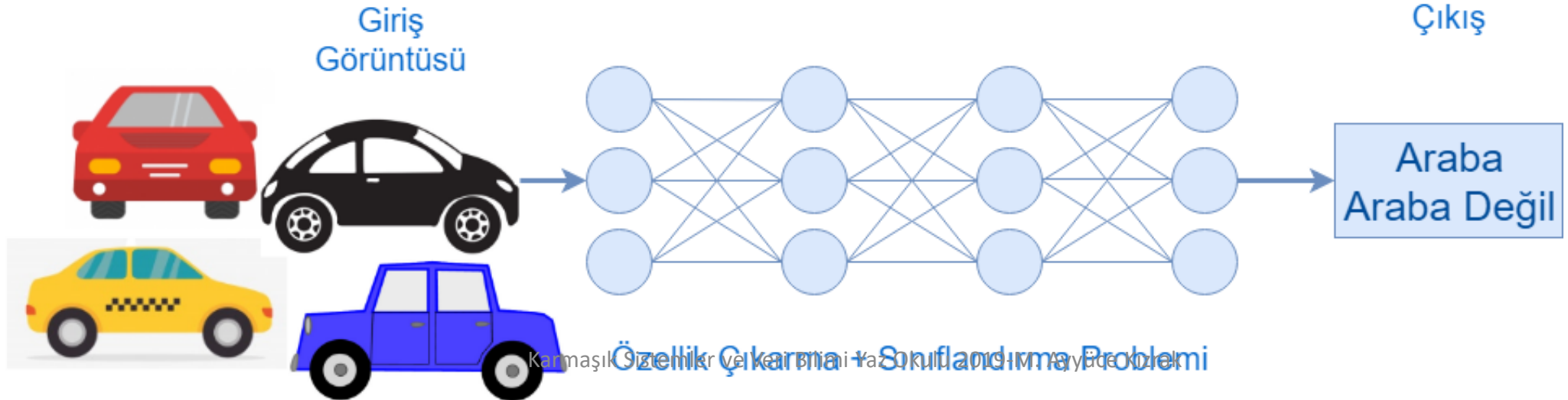
maliyet J



MAKİNE ÖĞRENMESİ



DERİN ÖĞRENME



Tinker With a **Neural Network** Right Here in Your Browser. Don't Worry, You Can't Break It. We Promise.



Epoch
000,000

Learning rate
0.03

Activation
Tanh

Regularization
None

Regularization rate
0

Problem type
Classification

DATA

Which dataset do you want to use?



Ratio of training to test data: 50%



Noise: 0



Batch size: 10

FEATURES

Which properties do you want to feed in?

- X_1
- X_2
- X_1^2
- X_2^2

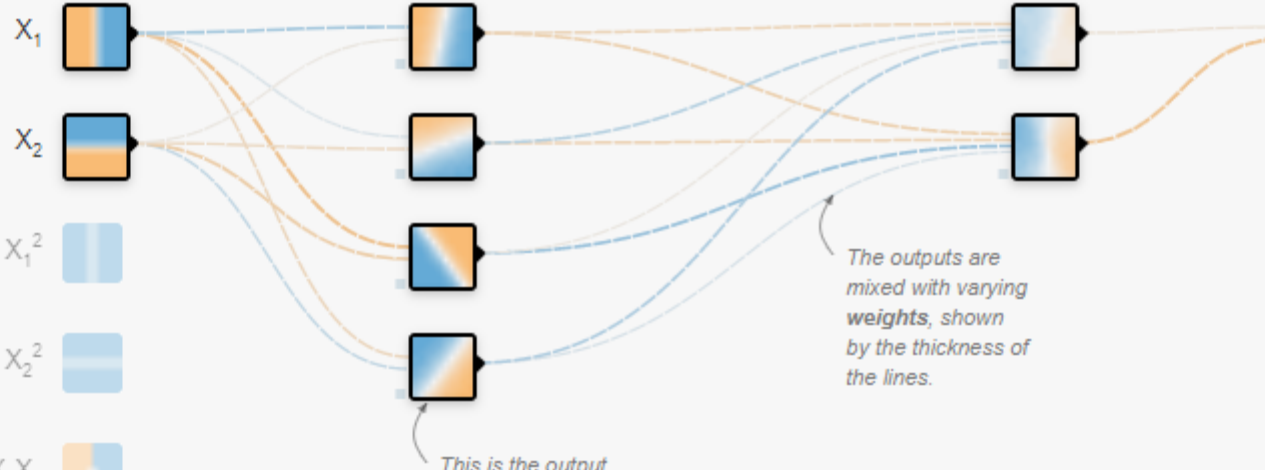
2 HIDDEN LAYERS

+ -

4 neurons

+ -

2 neurons



The outputs are mixed with varying **weights**, shown by the thickness of the lines.

OUTPUT

Test loss 0.560

Training loss 0.503

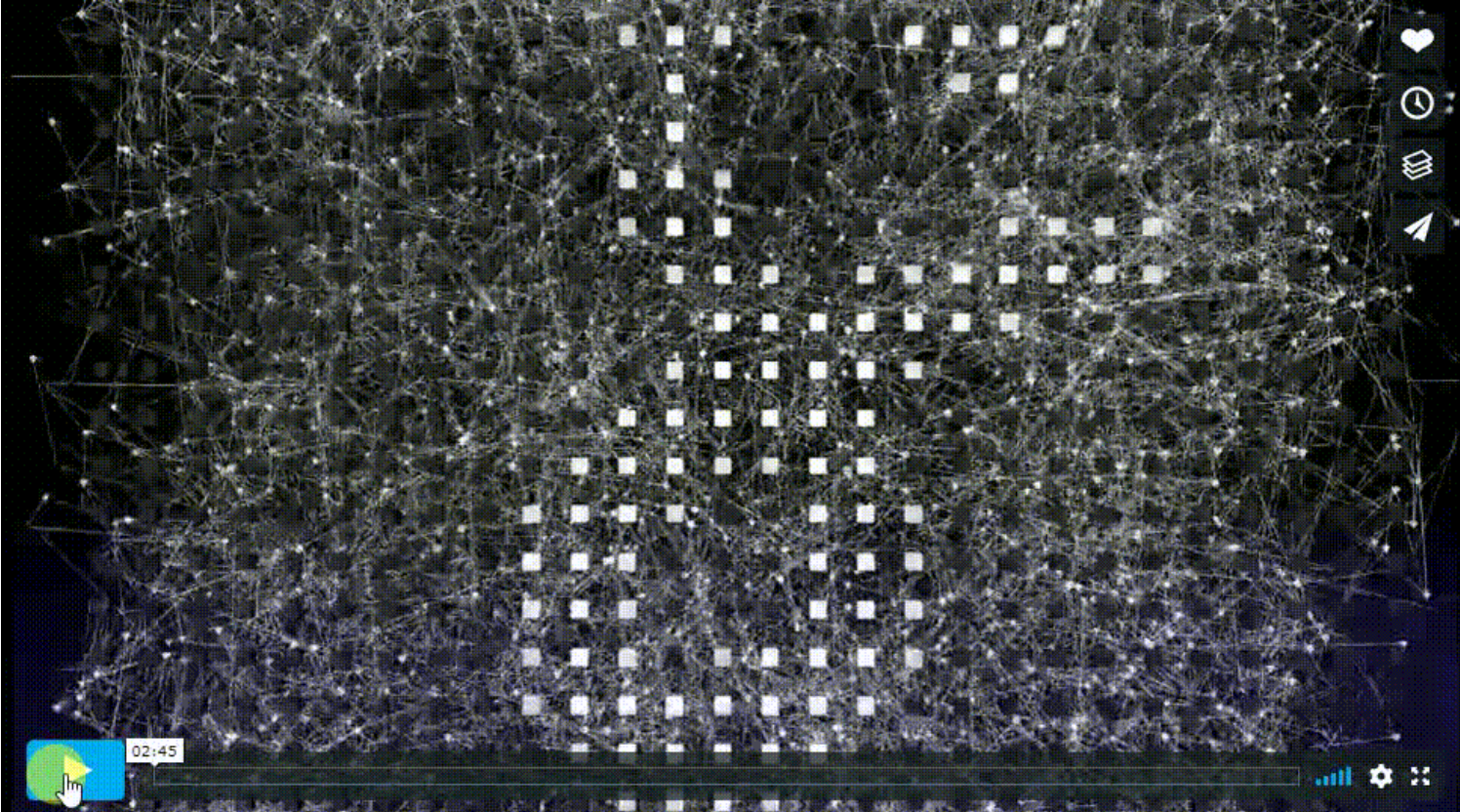


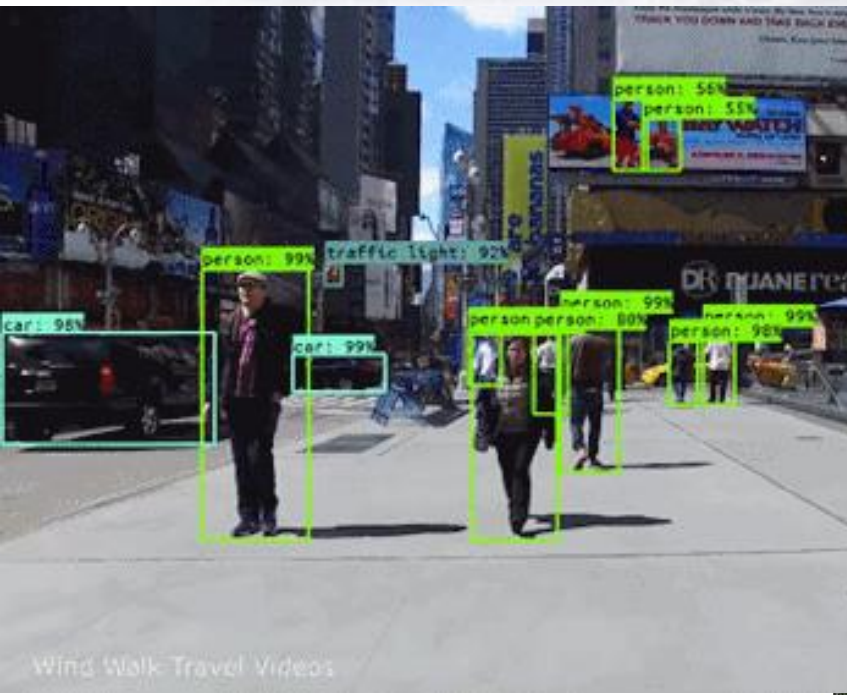
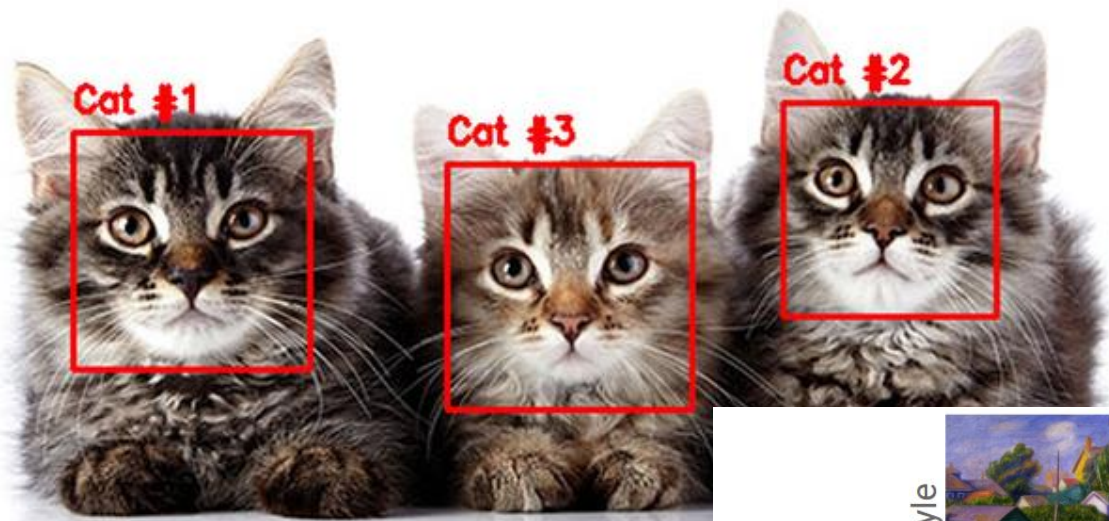
Görüntü Sınıflandırma

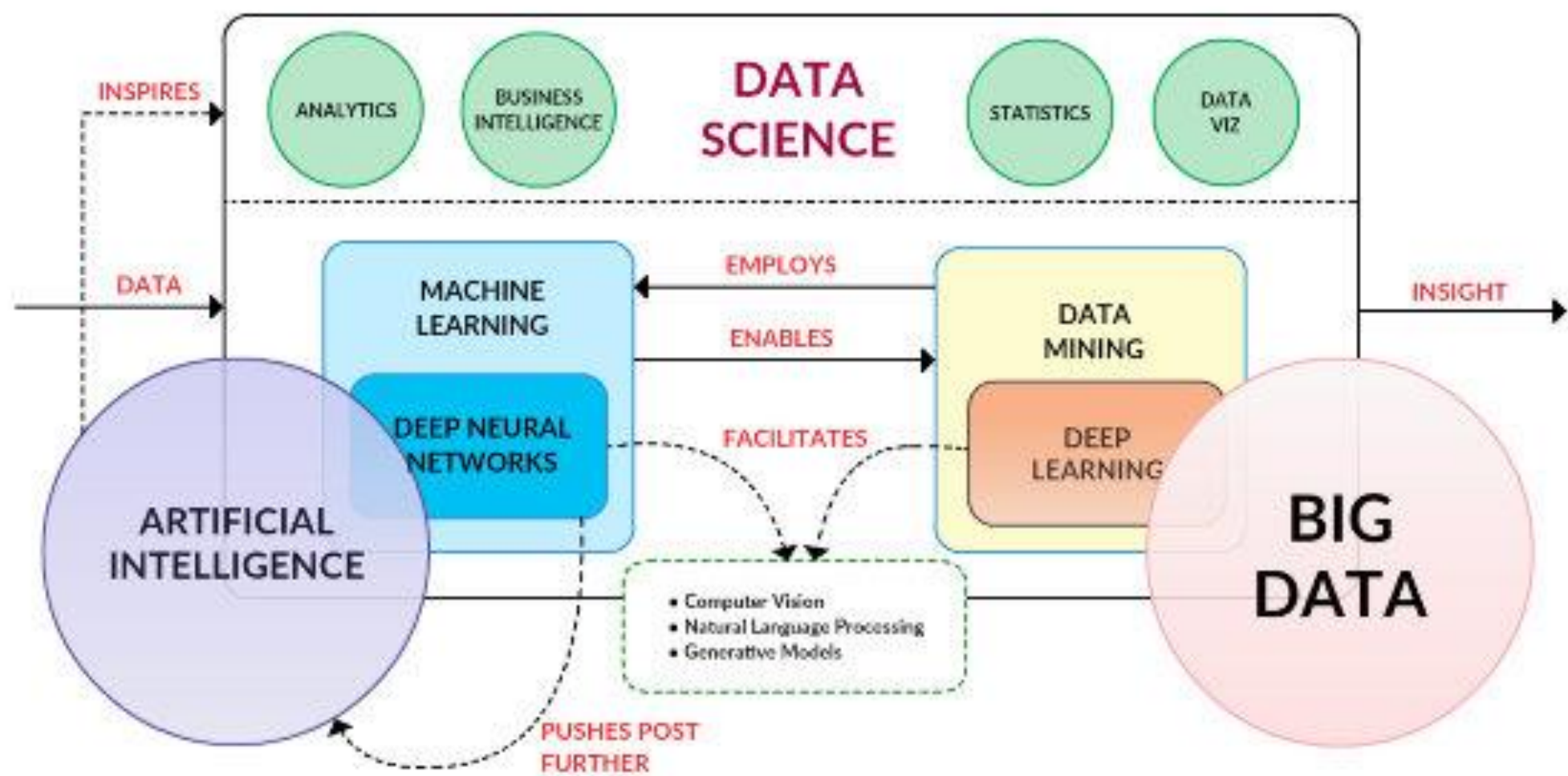


Köpek (1/0) ←

64x64x3









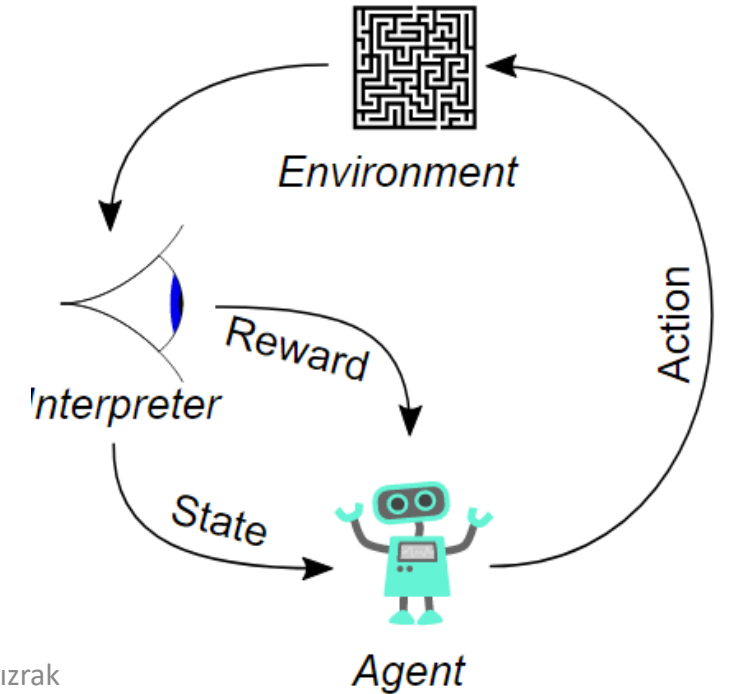
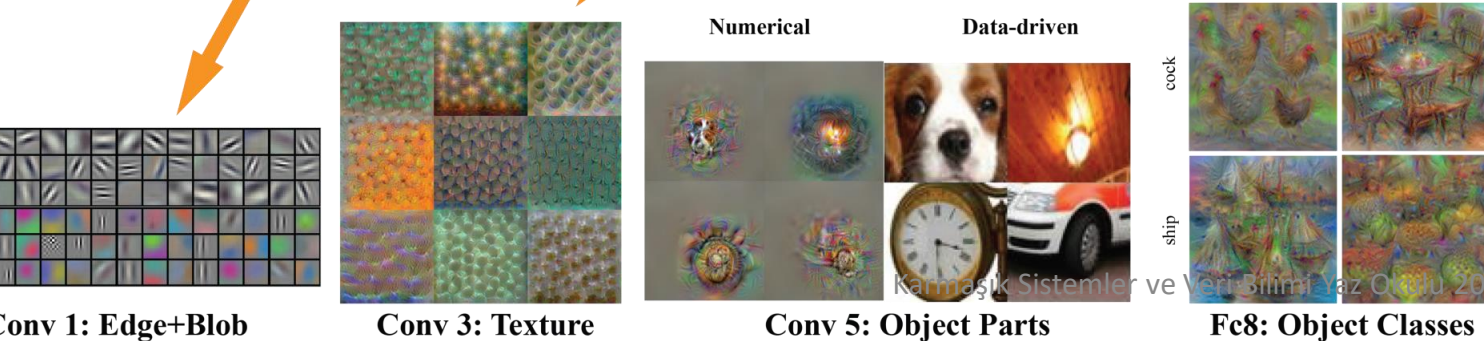
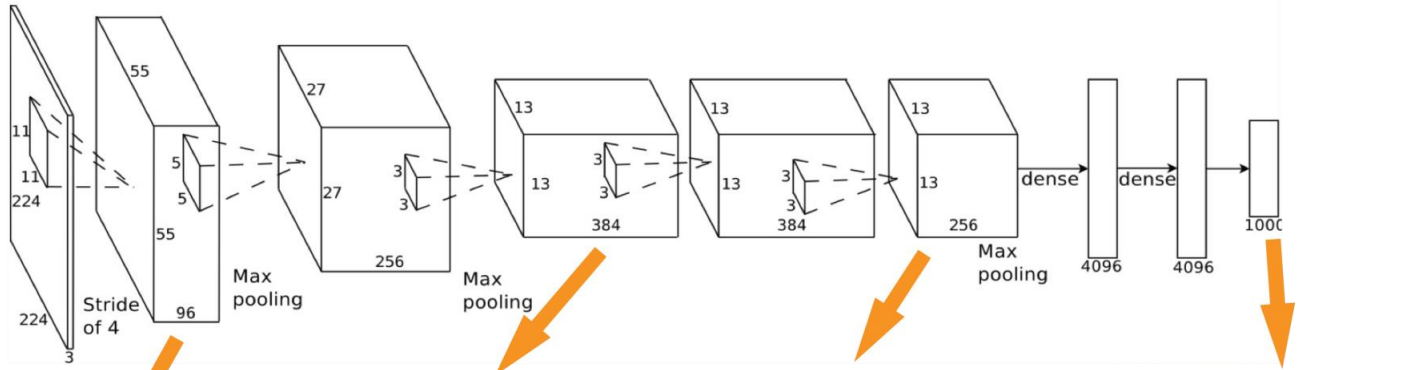
Data has a better idea

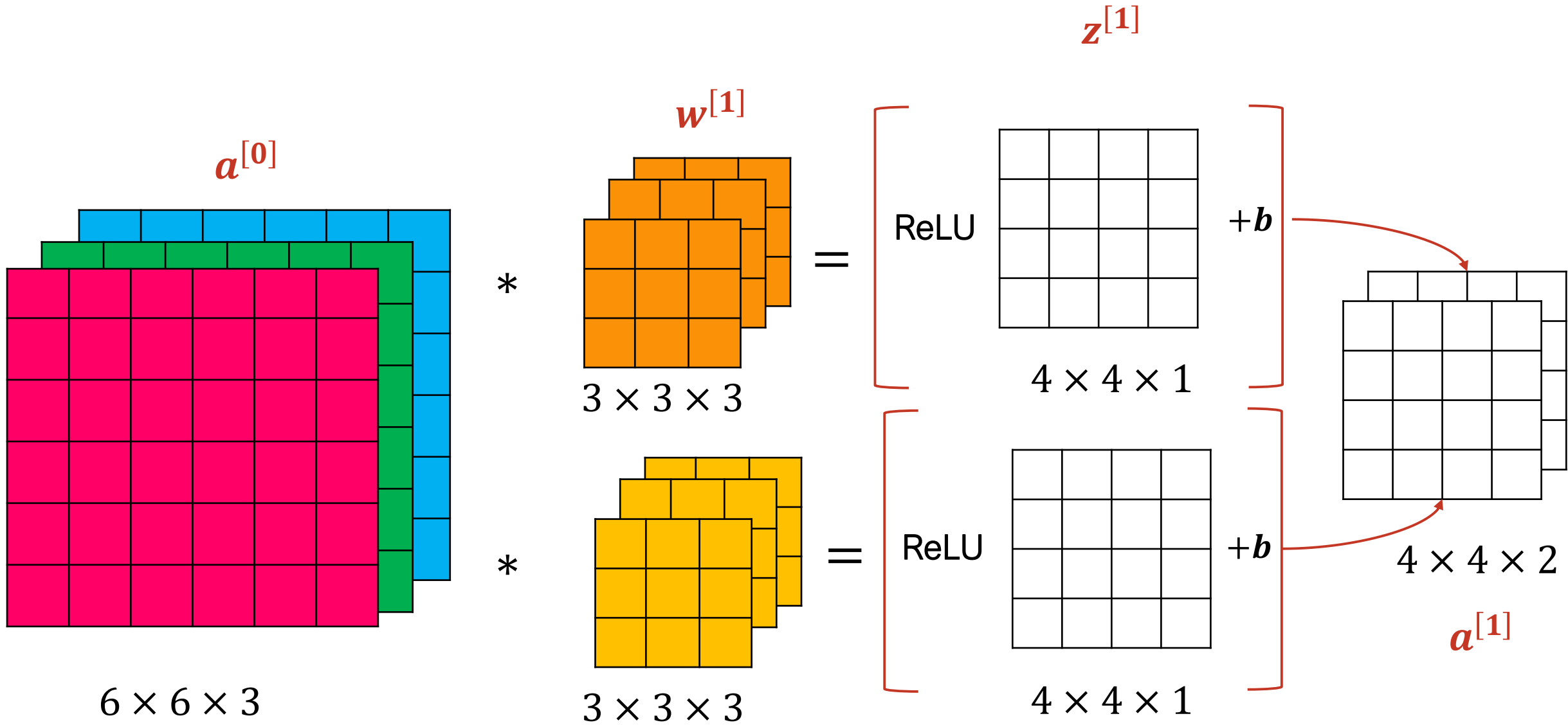
A person is captured mid-air, diving from a high, light-colored rocky cliff into the sea. The background features a vast, blue sky with scattered white clouds and a calm sea. The overall scene conveys a sense of adventure and depth.

DERİN ÖĞRENME MODELLERİ

Hangi problem için hangi modelleri tercih etmeliyim/öğrenmeliyim?

- **CNN** (Convolutional Neural Networks-Evrişimli Sinir Ağları): Nesne tanıma ve takip etme, stil transferi, kanser tespiti vb.
- **LSTM** (Long Short Term Memory-Uzun-Kısa Süreli Bellek): Doğal dil işleme, çeviri, chatbot, finans uygulamaları vb.
- **GAN** (Generative Adversarial Networks-Çekişmeli Üretici Ağlar): Sentetik veri üretme, sahte yüz üretme, stil transferi vb.
- **RL** (Reinforcement Learning-Pekiştirmeli Öğrenme): Kendi kendine ve az veriyle öğrenen yapay zeka sistemleri vb.





$$z^{[1]} = w^{[1]} \cdot a^{[0]} + b^{[1]}$$

$$a^{[1]} = g^{[1]}(z^{[1]})$$

Filtre boyutu : $f^{[l]}$
Piksel doldurma : $p^{[l]}$
Adım kaydırma : $s^{[l]}$
Filtre (kanal) sayısı : $n_C^{[l]}$

Her bir filtrenin boyutu : $f^{[l]} \times f^{[l]} \times n_C^{[l-1]}$

Aktivasyon fonksiyonu : $a^{[l]} \rightarrow n_H^{[l]} \times n_W^{[l]} \times n_C^{[l]}$

Ağırlıklar : $f^{[l]} \times f^{[l]} \times n_C^{[l-1]} \times n_C^{[l]}$

Bias : $n_C^{[l]} \rightarrow (1, 1, 1, n_C^{[l]})$

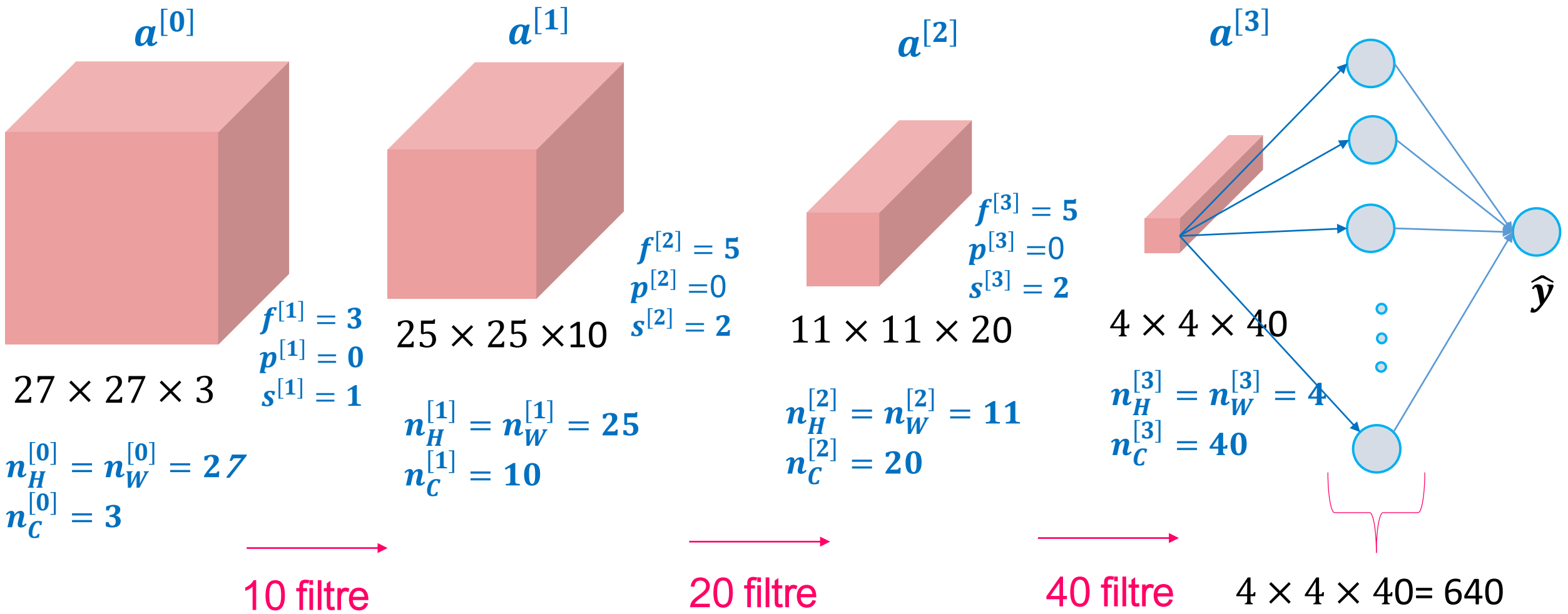
Giriş : $n_H^{[l-1]} \times n_W^{[l-1]} \times n_C^{[l-1]}$

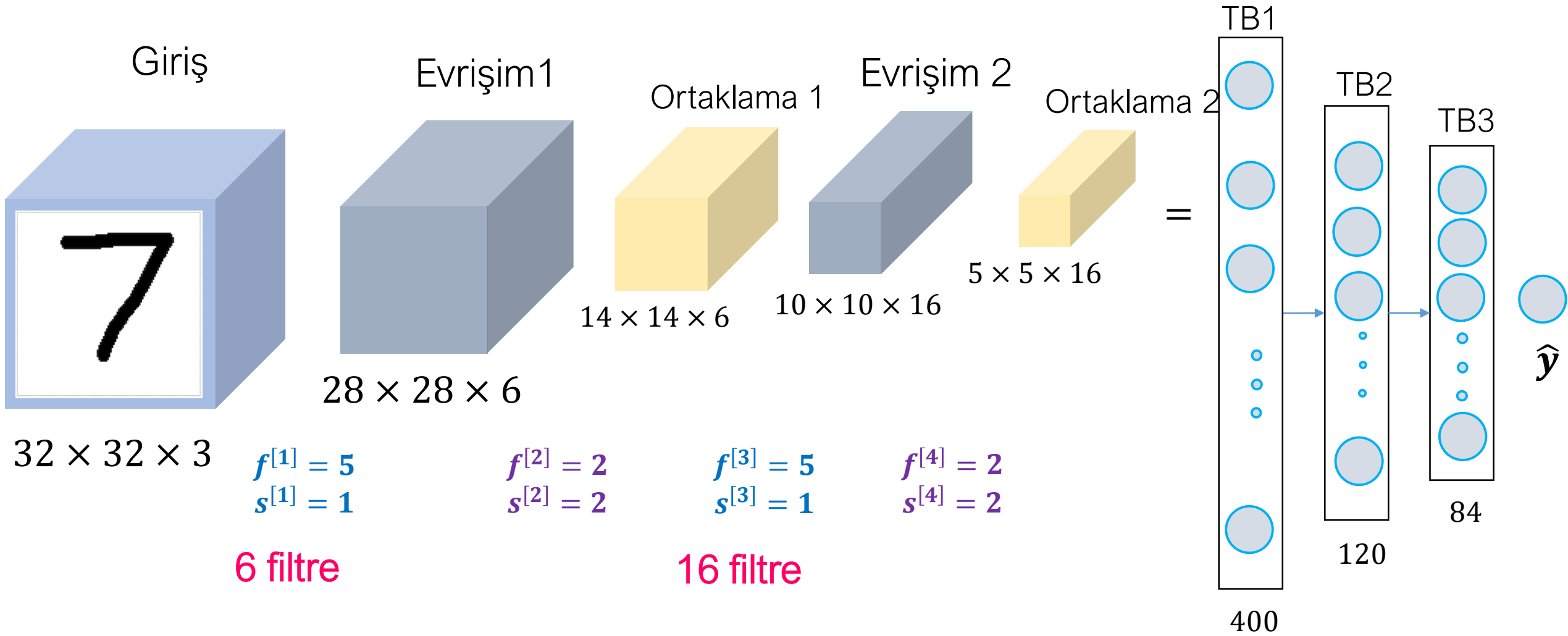
Çıkış : $n_H^{[l]} \times n_W^{[l]} \times n_C^{[l]}$

$$n_{H,W}^{[l]} : \left[\frac{n_H^{[l-1]} + 2p - f^{[l]}}{s^{[l]}} + 1 \right]$$

Bir Evrişimli Sinir Ağı için Gereken Katmanlar

- Evrişim Katmanı (Aktivasyon fonksiyonu, Bias değeri)
- Ortaklama Katmanı (Maksimum ya da ortalama ortaklama)
- Tam/Tüm Bağlantı Katmanı (Klasik yapay sinir ağı bağlantıları)





Çıkışta 10 sınıflı rakam tanıma yapılacaksa: softmax(10)

$n_H, n_W \downarrow$
 $n_C \uparrow$

LeNet 5 – 1980'lerden - 1998

PROC. OF THE IEEE, NOVEMBER 1998

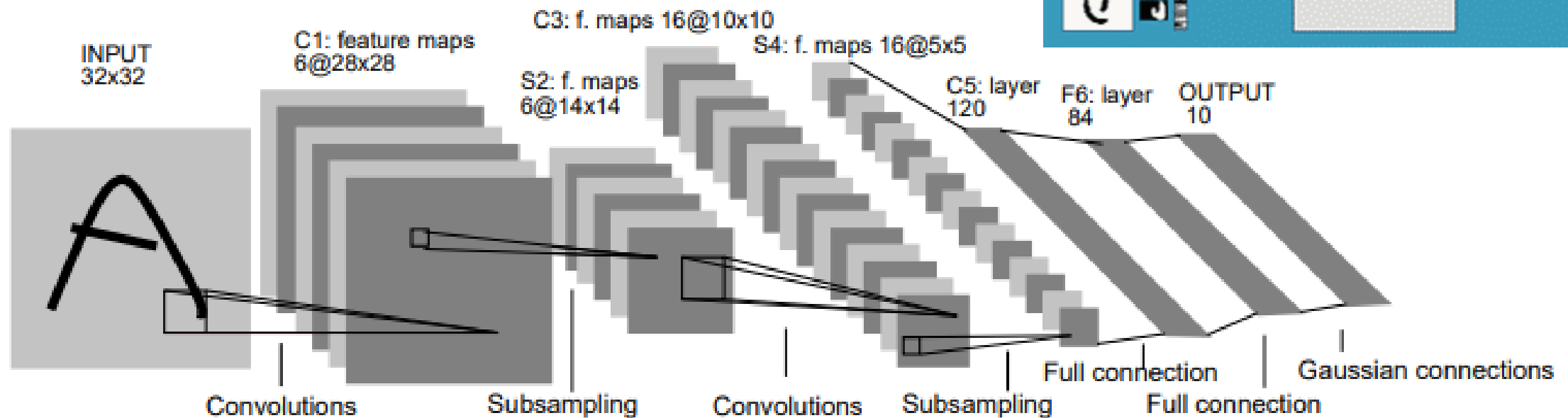
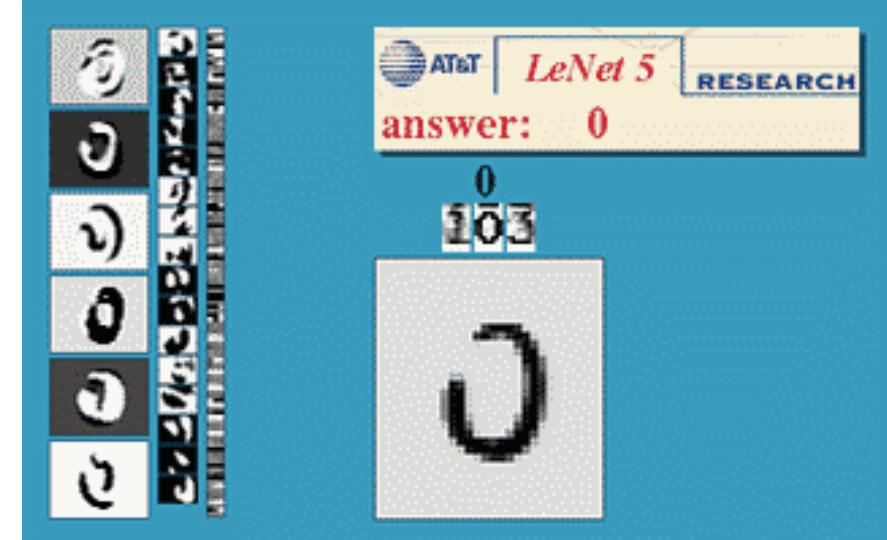
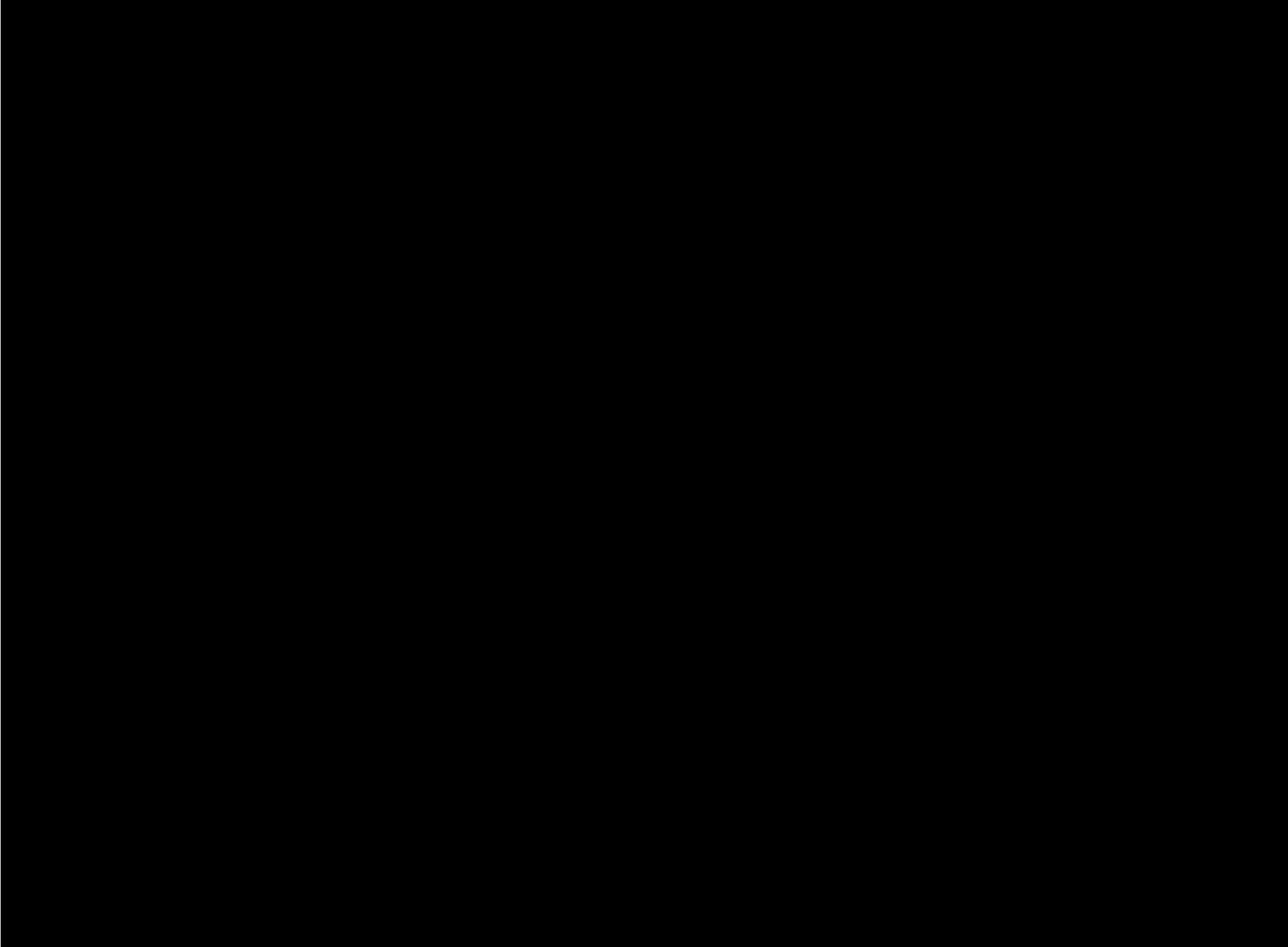
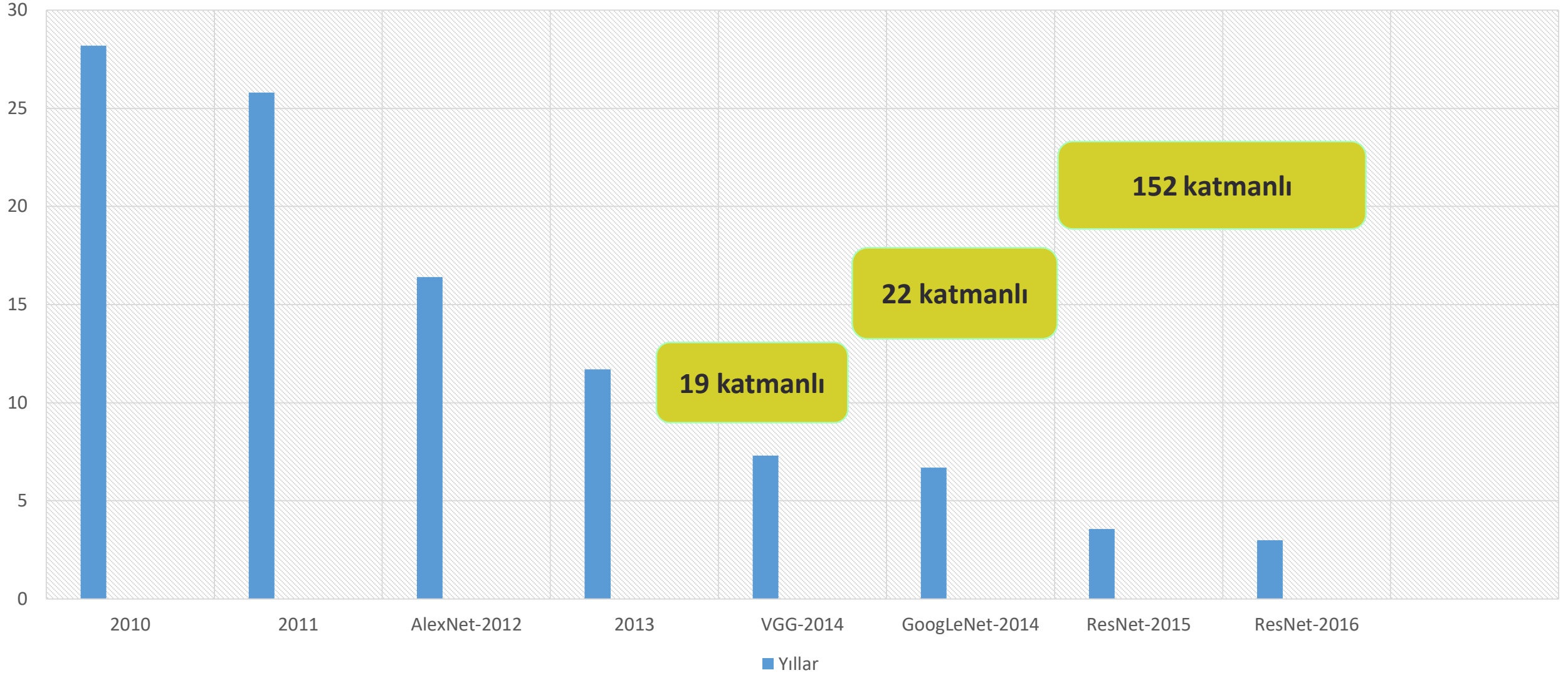
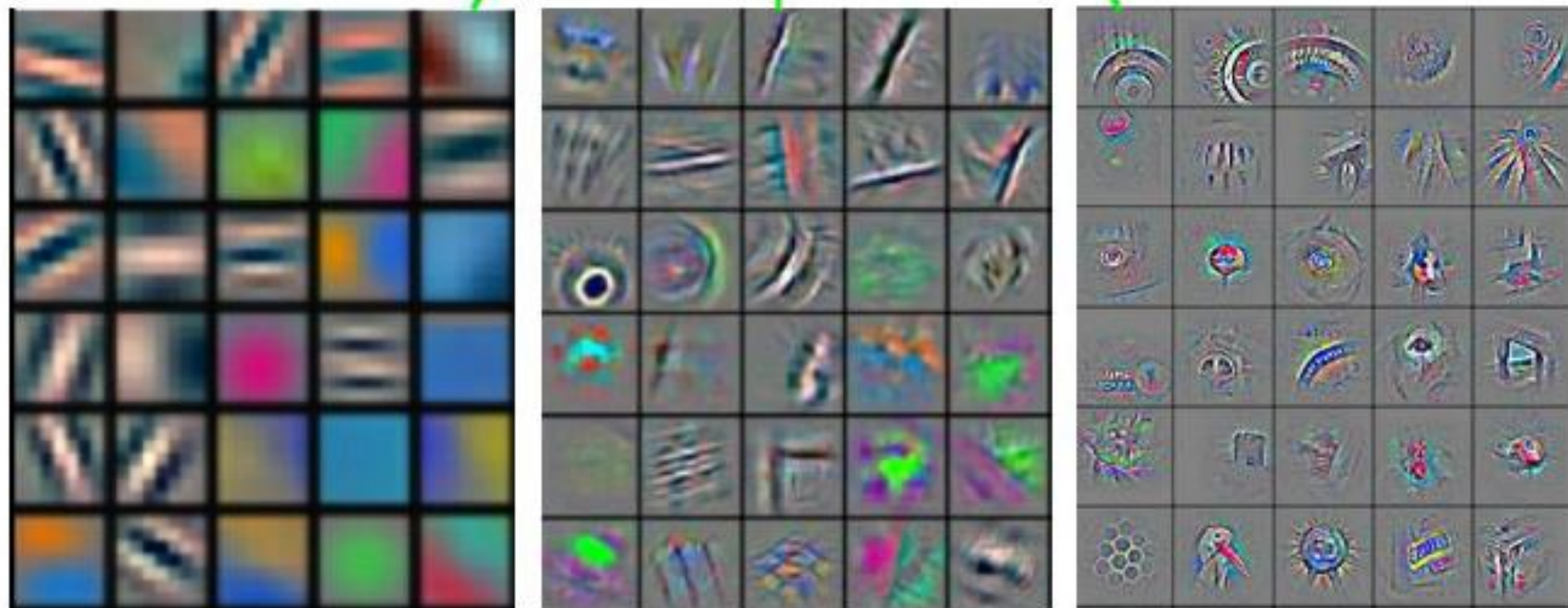


Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.



EĞİTİM HATASI (%)





Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

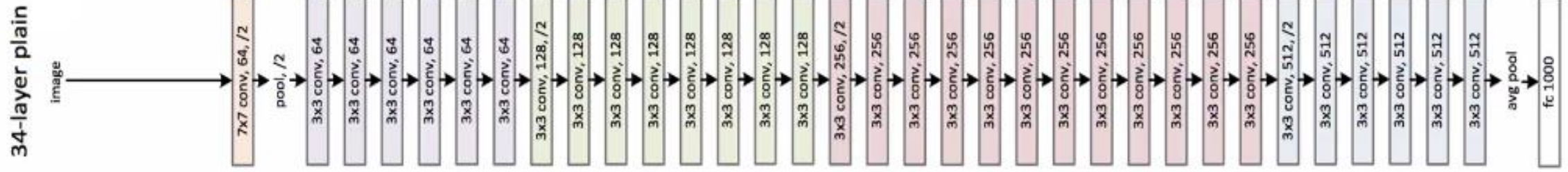


DAHA DERİNE

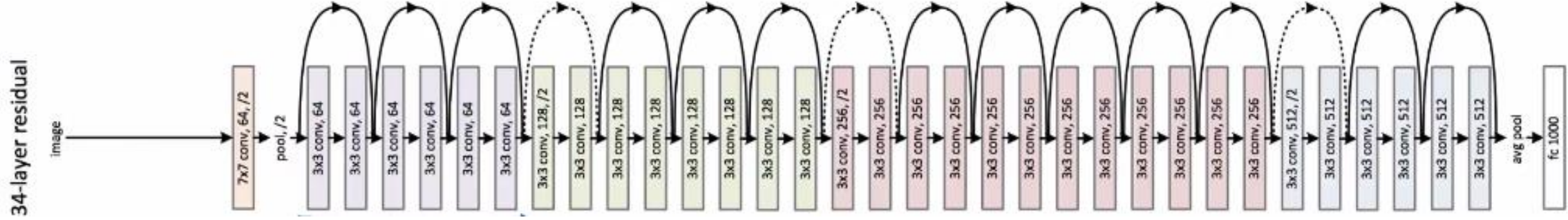


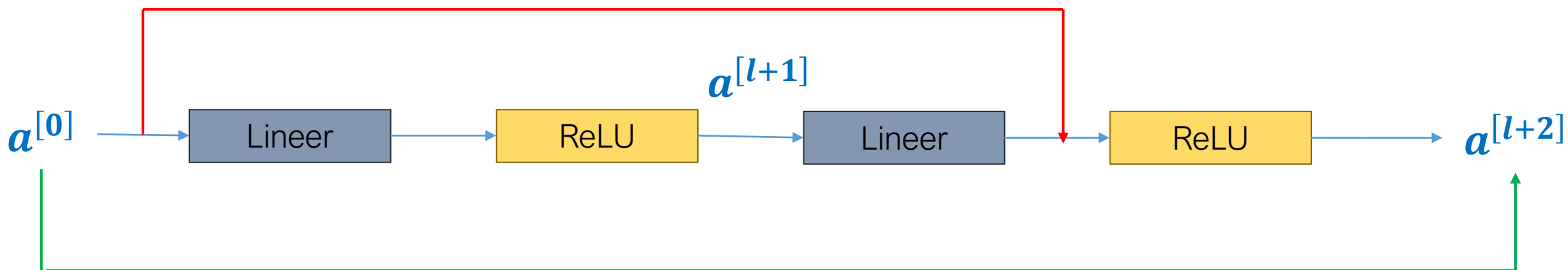
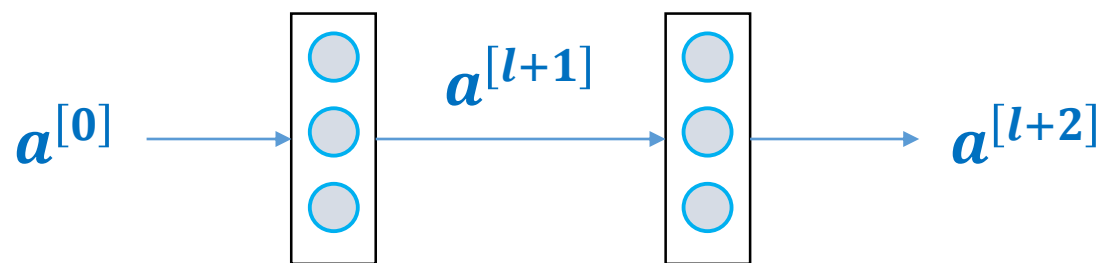
ResNets 152

Plain



ResNet





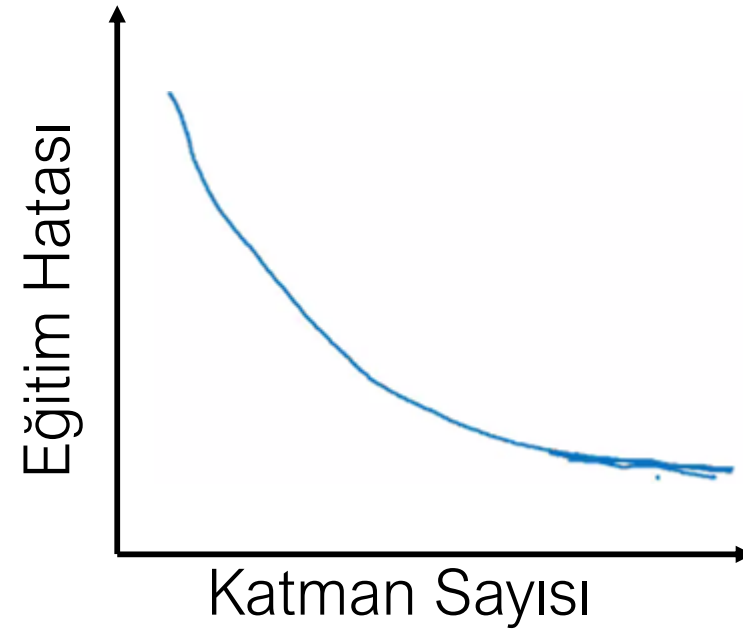
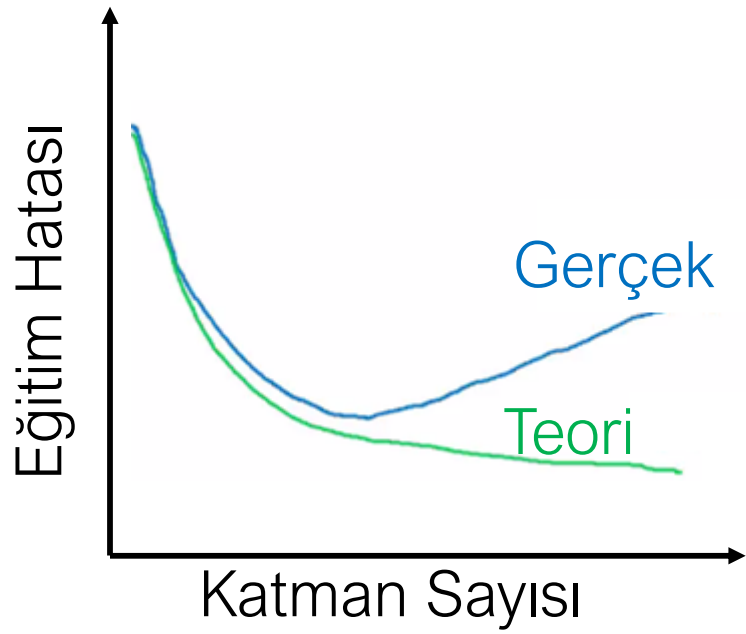
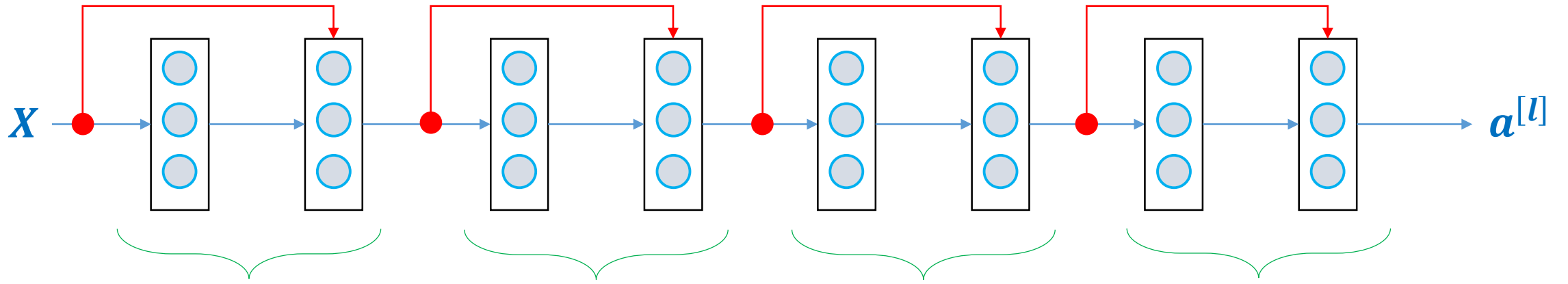
$$z^{[l+1]} = W^{[l+1]} a^l + b^{[l+1]}$$

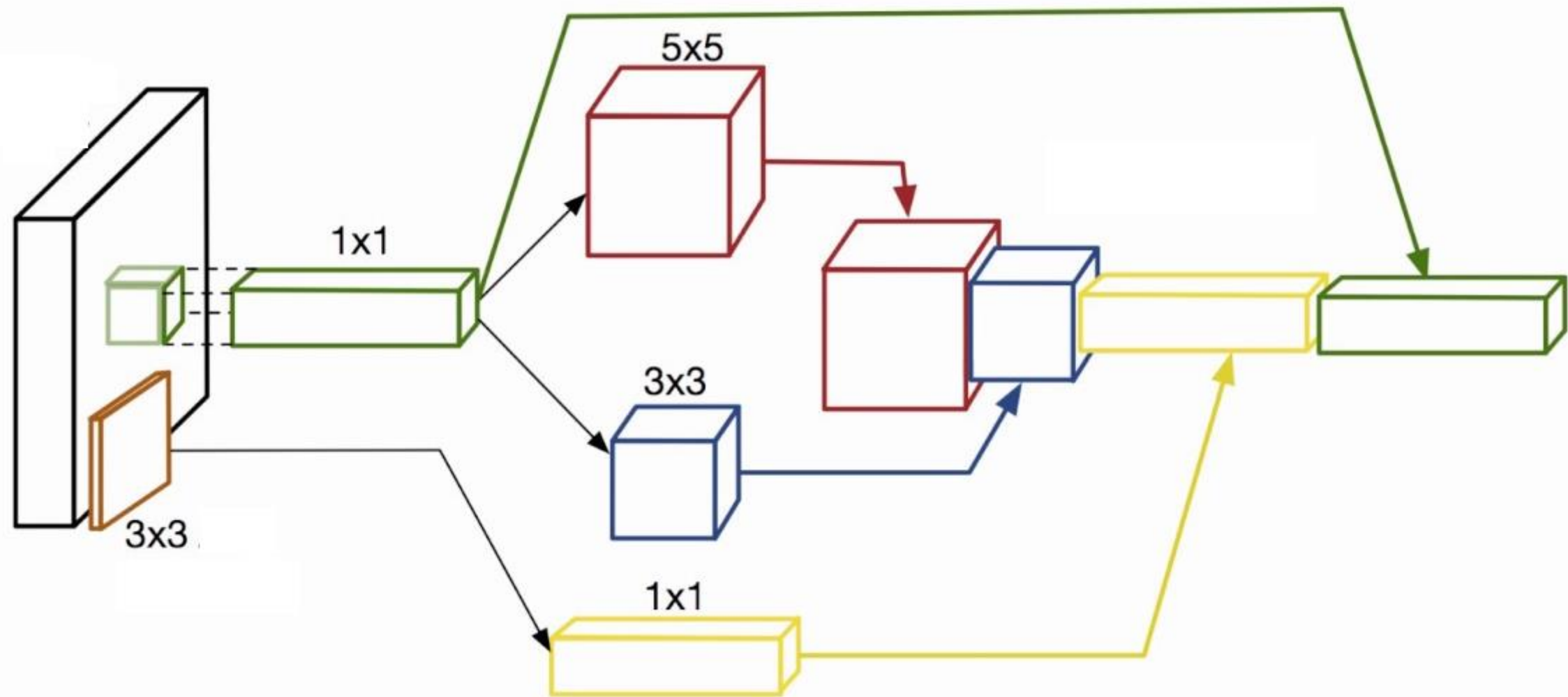
$$a^{[l+1]} = g(z^{[l+1]})$$

$$z^{[l+2]} = W^{[l+2]} a^{[l+1]} + b^{[l+2]}$$

~~$$a^{[l+2]} = g(z^{[l+2]})$$~~

$$a^{[l+2]} = g(z^{[l+2]} + \boxed{a^{[l]}})$$



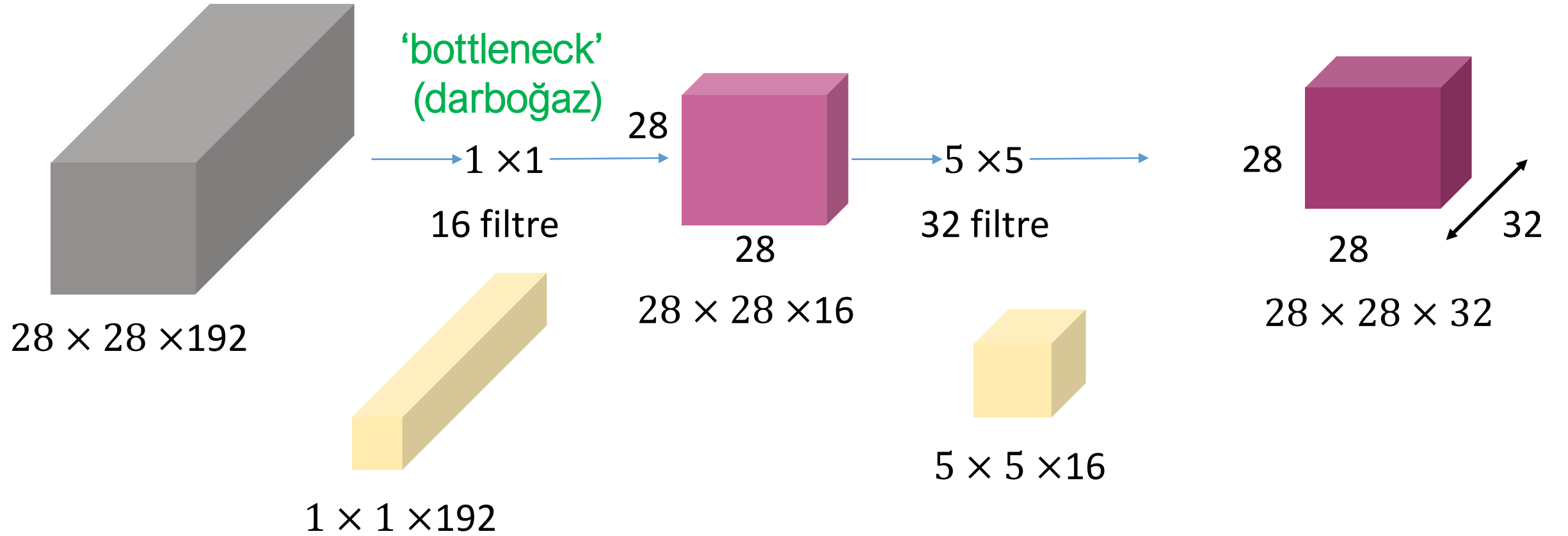


Bu koşulda; 1x1 evrişim katmanında: $(28 \times 28 \times 16) \times (1 \times 1 \times 192) = 2,4$ milyon parametre

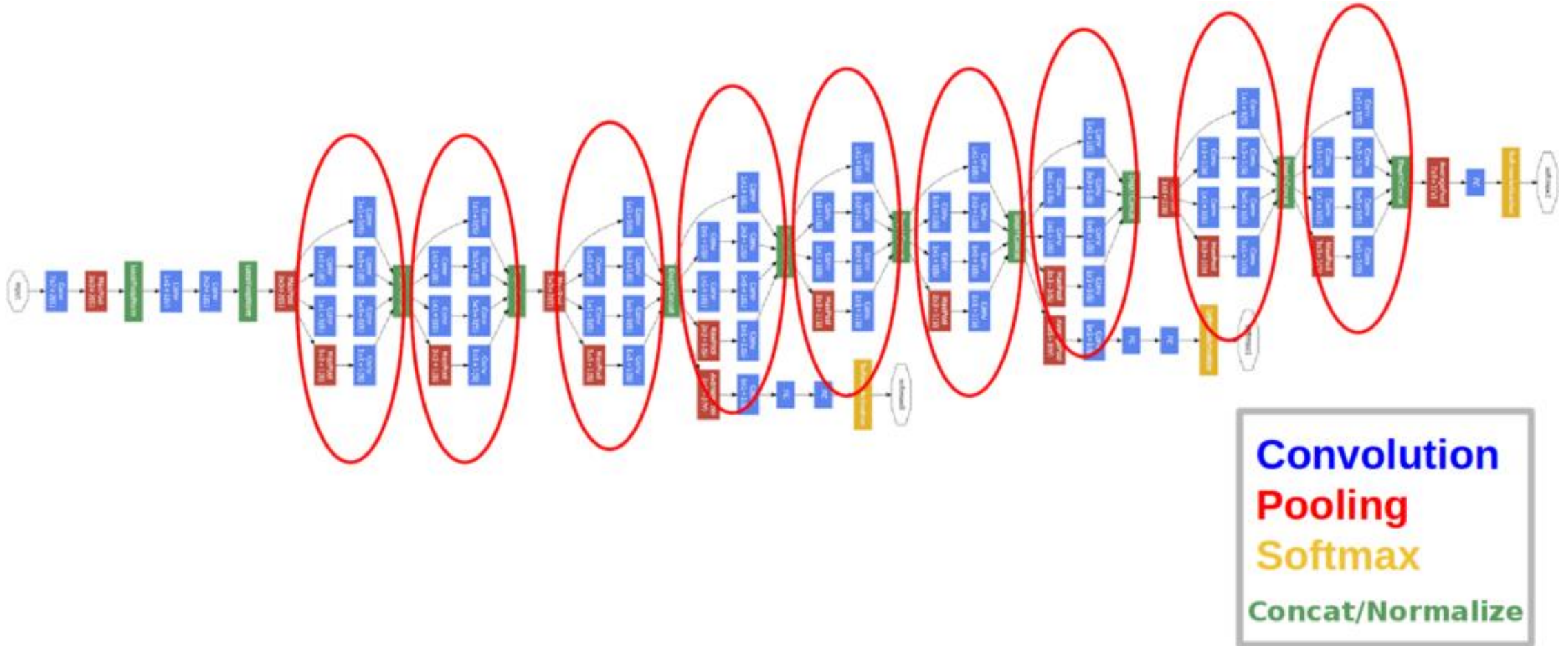
5x5 evrişim katmanında: $(28 \times 28 \times 32) \times (5 \times 5 \times 16) = 10$ milyon parametre

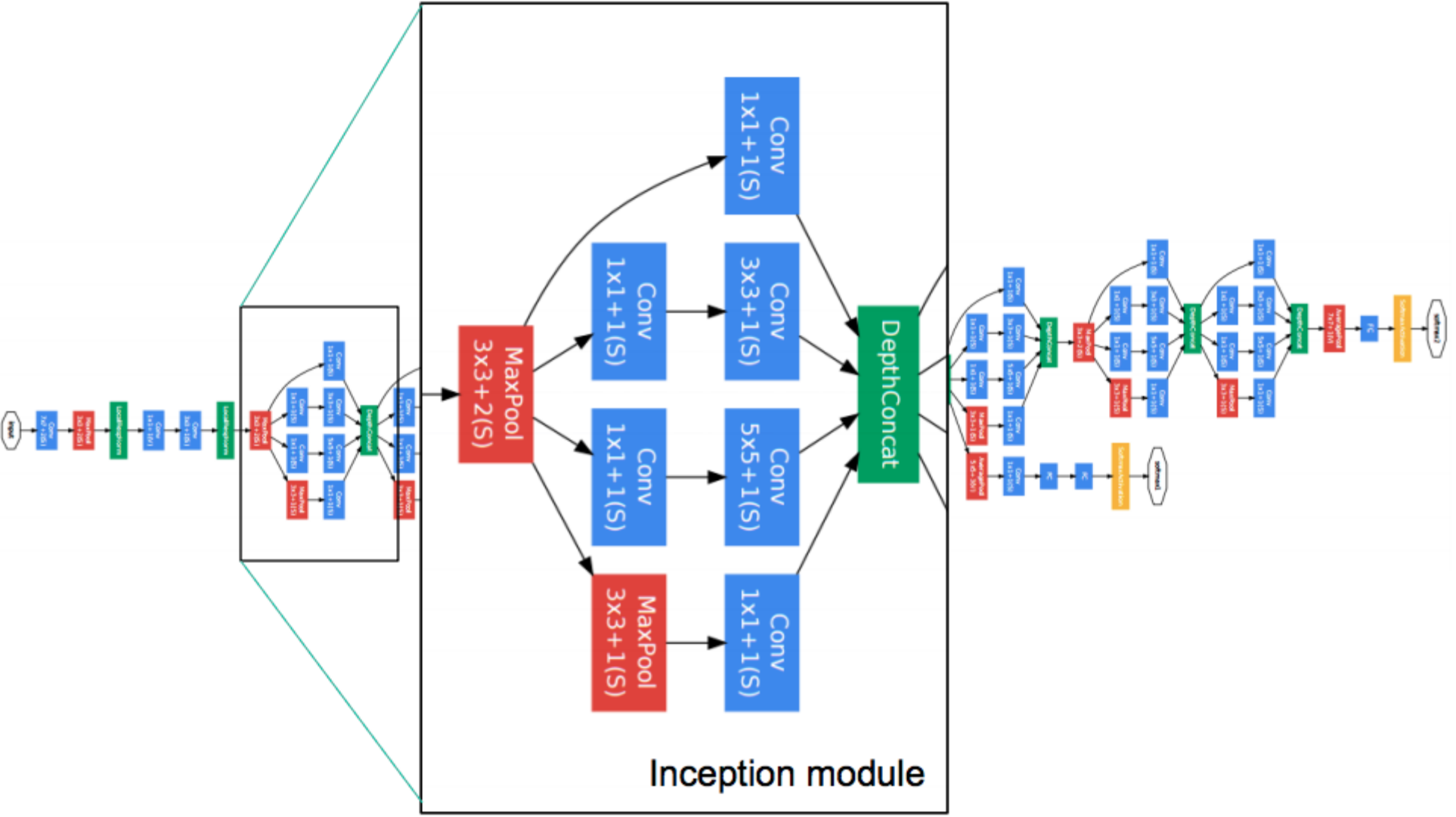
Toplamda 12.4 milyon parametre

İlk duruma göre yaklaşık 10 kat daha az parametre hesabı son derece çarpıcıdır.



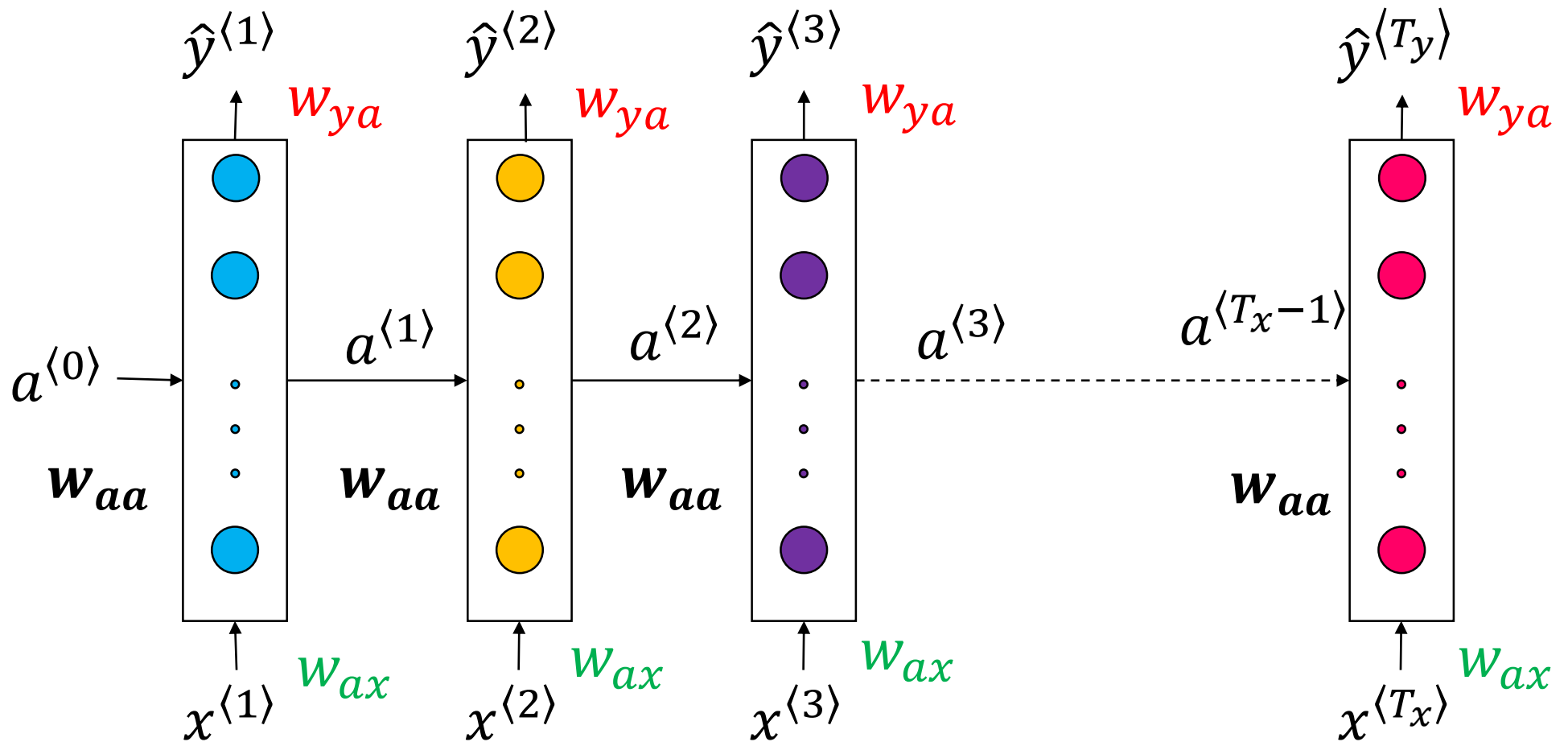
GoogLeNet - 2014





ÖZYİNELEMELİ SİNİR AĞLARI





$$a^{(0)} = \vec{0}$$

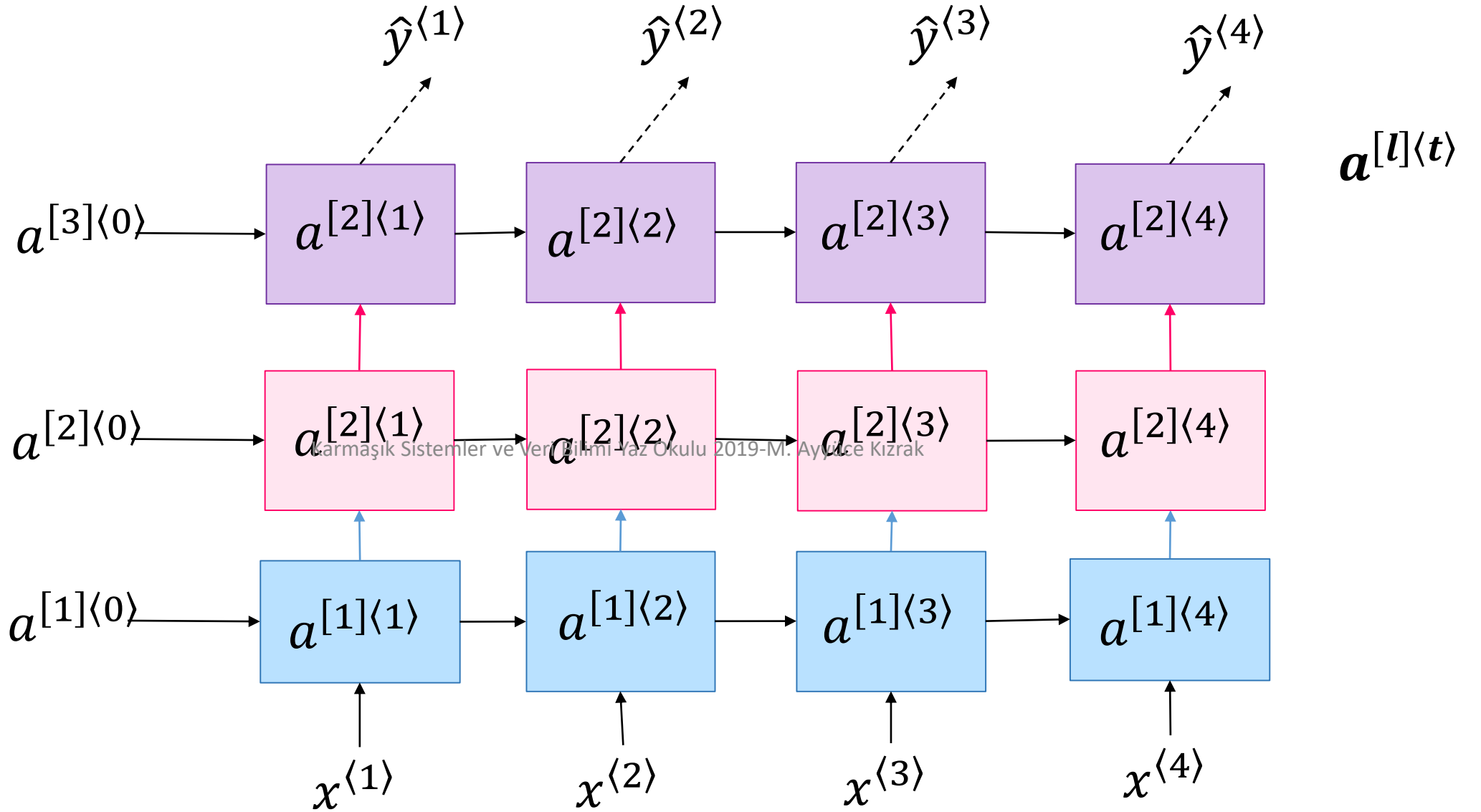
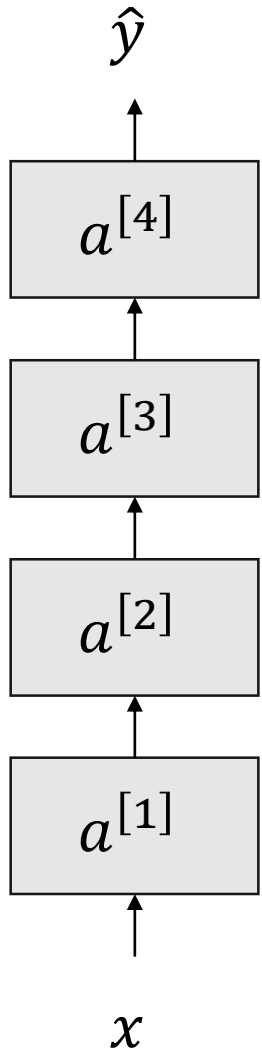
$$a^{(1)} = g_1(W_{aa}a^{(0)} + W_{ax}x^{(1)} + b_a)$$

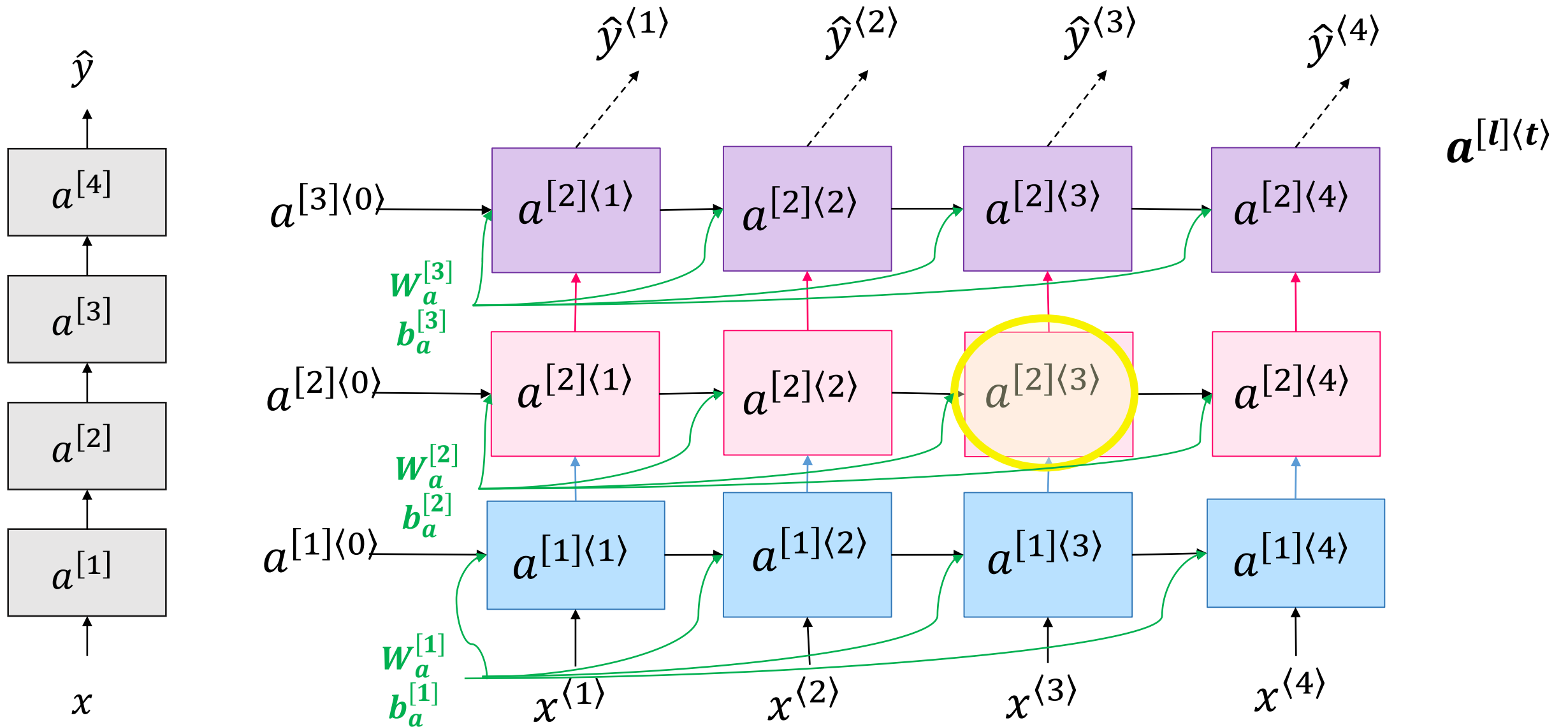
$$\hat{y}^{(1)} = g_2(W_{ya}a^{(1)} + b_y)$$

$$a^{(t)} = g(W_{aa}a^{(t-1)} + W_{ax}x^{(t)} + b_a) \quad \text{Tanh/ReLU}$$

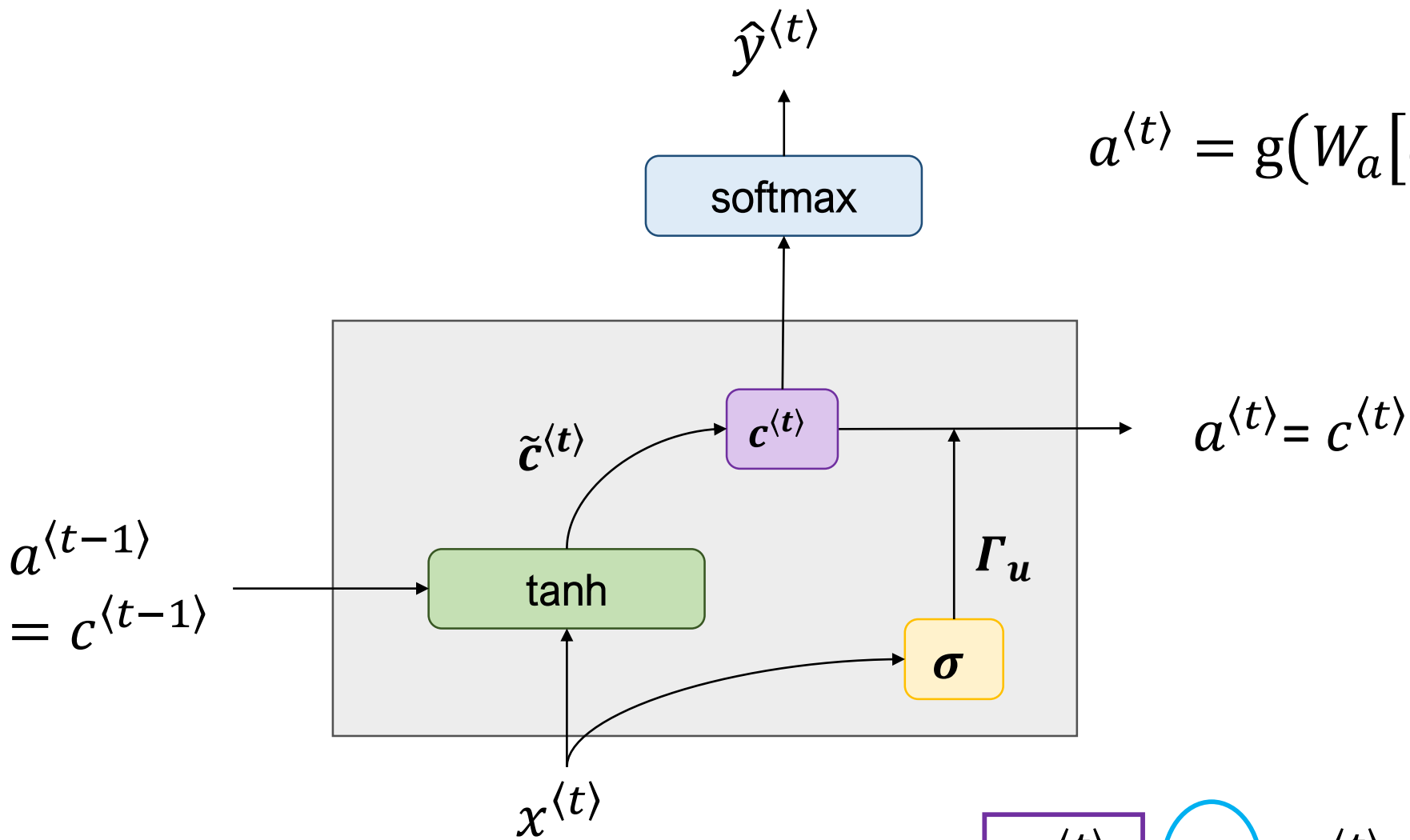
$$\hat{y}^{(t)} = g(W_{ya}a^{(t)} + b_y) \quad \text{Sigmoid}$$

Karmaşık Sistemler ve Veri Bilimi Yaz Okulu 2019-M. Ayyüce Kızrak





$$a^{[2]\langle 3 \rangle} = g \left(W_a^{[2]} [a^{[2]\langle 2 \rangle}, a^{[1]\langle 3 \rangle}] + b_a \right)$$



$$a^{t} = g(W_a[a^{t-1}, x^{t}] + b_a)$$

$$a^{t} = c^{t}$$

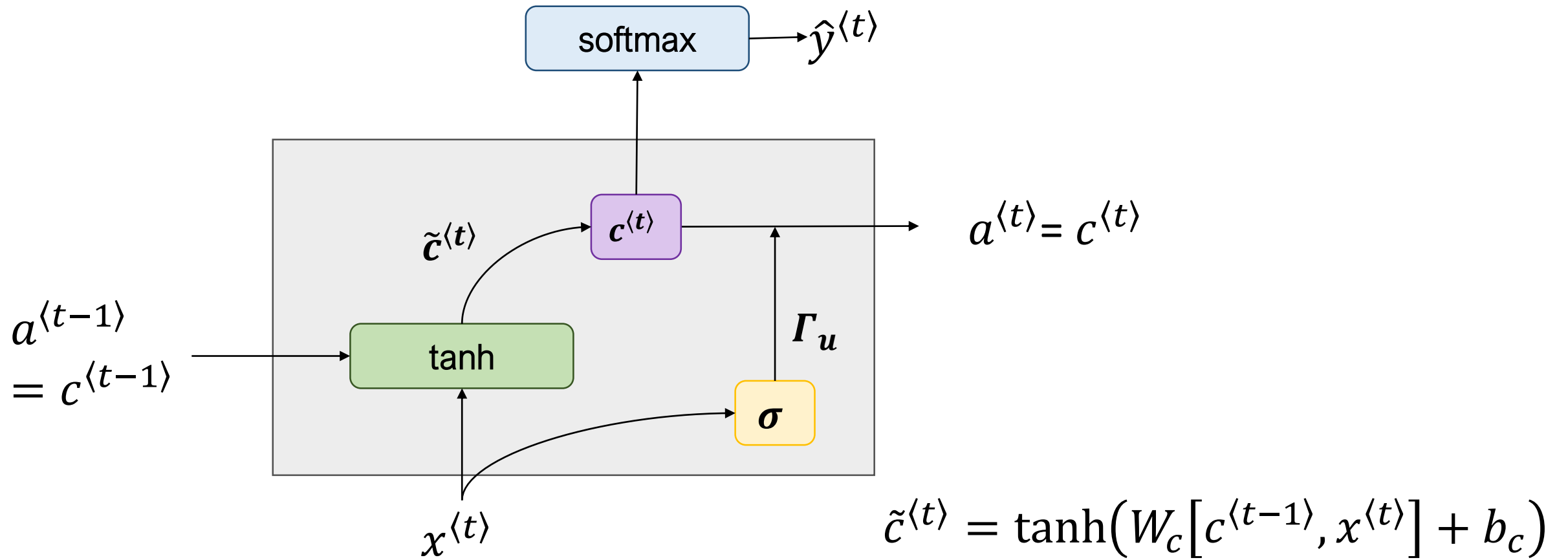
$$\tilde{c}^{t} = \tanh(W_c[c^{t-1}, x^{t}] + b_c)$$

$$c^{t} = \Gamma_u * \tilde{c}^{t} + (1 - \Gamma_u) * c^{t-1}$$

$$\Gamma_u = \sigma(W_u[c^{t-1}, x^{t}] + b_u)$$

$$\Gamma_u = 1 \quad \Gamma_u = 0 \quad \Gamma_u = 0 \quad \Gamma_u = 0 \quad \Gamma_u = 1$$

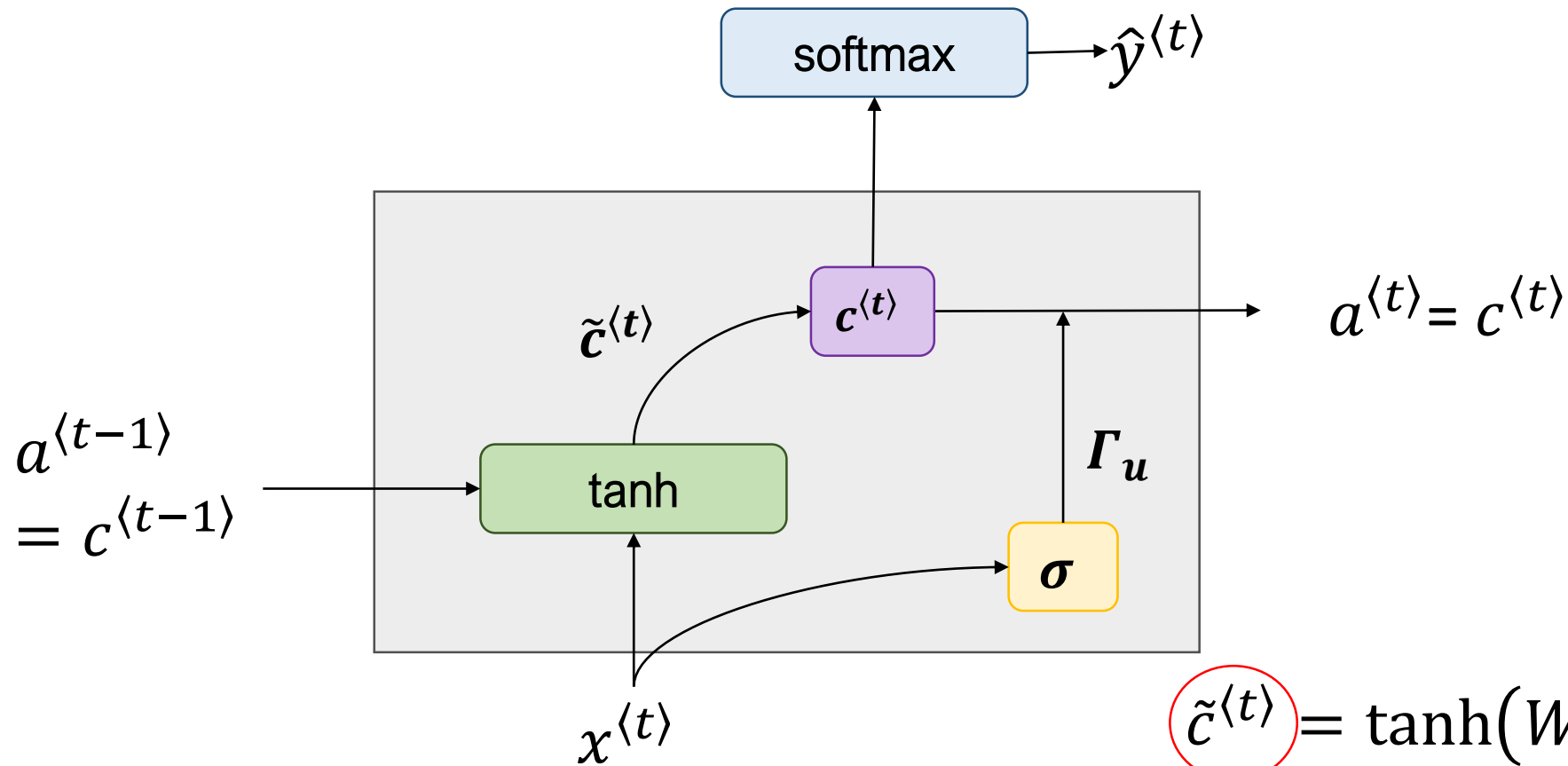
The dogs which already gone were so cute.



$$\tilde{c}^{t} = \tanh(W_c [c^{t-1}, x^{t}] + b_c)$$

Güncelleme Geçiti $\Gamma_u = \sigma(W_u [c^{t-1}, x^{t}] + b_u)$

$$c^{t} = \Gamma_u * \tilde{c}^{t} + (1 - \Gamma_u) * c^{t-1}$$



$$\tilde{c}^{(t)} = \tanh(W_c[\Gamma_r * c^{(t-1)}, x^{(t)}] + b_c)$$

Güncelleme Geçiti $\Gamma_u = \sigma(W_u[c^{(t-1)}, x^{(t)}] + b_u)$

İlgililik Geçiti $\Gamma_r = \sigma(W_r[c^{(t-1)}, x^{(t)}] + b_r)$

$$c^{(t)} = \Gamma_u * \tilde{c}^{(t)} + (1 - \Gamma_u) * c^{(t-1)}$$

$$\tilde{c}^{\langle t \rangle} = \tanh(W_c[\Gamma_r * c^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_c) \longrightarrow \tilde{c}^{\langle t \rangle} = \tanh(W_c[a^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_c)$$

Güncelleme Geçiti

$$\Gamma_u = \sigma(W_u[c^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_u) \longrightarrow \Gamma_u = \sigma(W_u[a^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_u)$$

İlgililik Geçiti

$$\Gamma_r = \sigma(W_r[c^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_r)$$

Unutma Geçiti

$$\Gamma_f = \sigma(W_f[a^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_r)$$

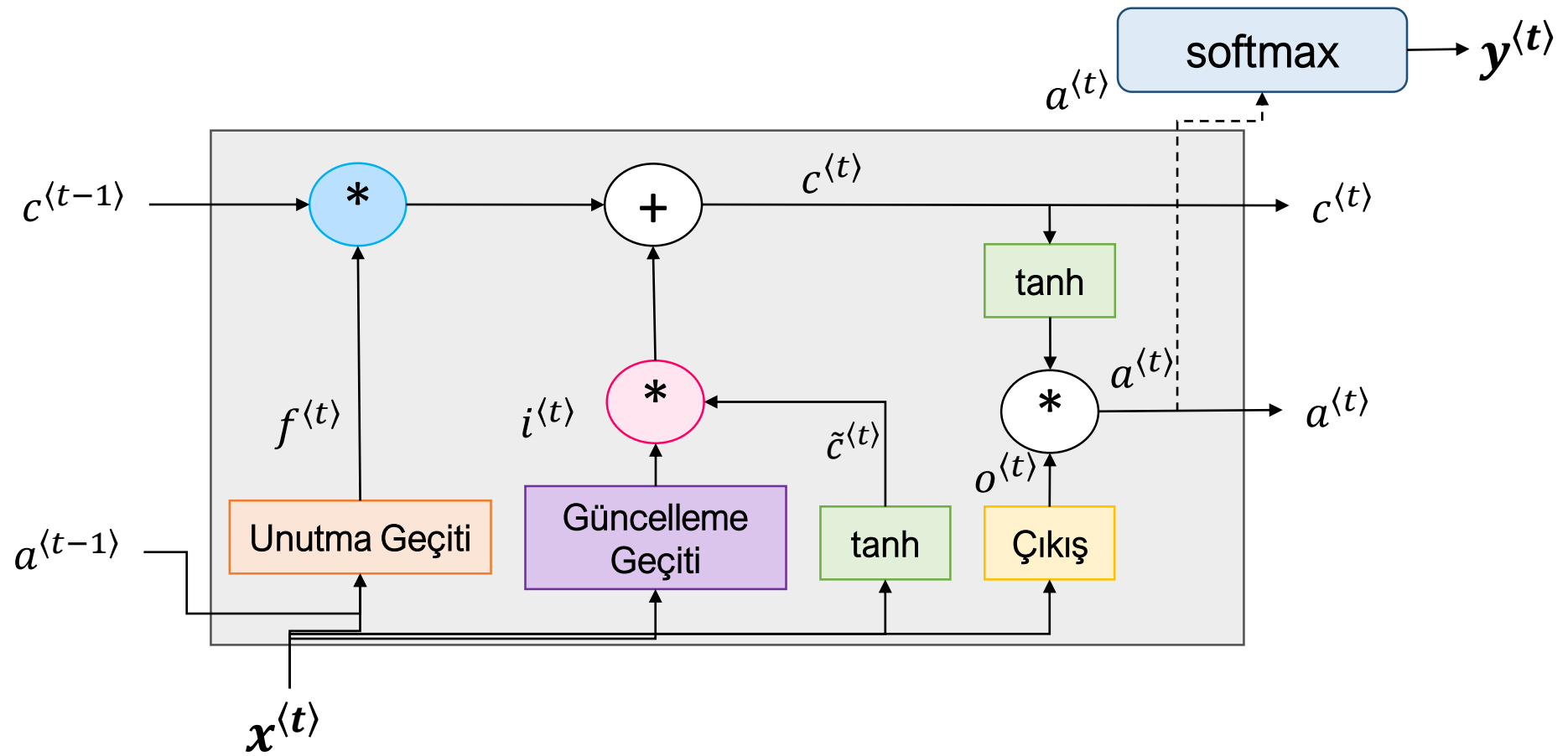
Çıkış

$$\Gamma_o = \sigma(W_o[a^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_o)$$

$$c^{\langle t \rangle} = \Gamma_u * \tilde{c}^{\langle t \rangle} + (1 - \Gamma_u) * c^{\langle t-1 \rangle} \longrightarrow c^{\langle t \rangle} = \Gamma_u * \tilde{c}^{\langle t \rangle} + (\Gamma_f) * c^{\langle t-1 \rangle}$$

$$c^{\langle t \rangle} = a^{\langle t \rangle}$$

$$\Gamma_o * c^{\langle t \rangle} = a^{\langle t \rangle}$$



$$\tilde{c}^{(t)} = \mathbf{tanh}(W_c[a^{(t-1)}, x^{(t)}] + b_c)$$

$$\Gamma_u = \sigma(W_u[a^{(t-1)}, x^{(t)}] + b_u)$$

$$\Gamma_f = \sigma(W_f[a^{(t-1)}, x^{(t)}] + b_r)$$

$$\Gamma_o = \sigma(W_o[a^{(t-1)}, x^{(t)}] + b_o)$$

$$c^{(t)} = \Gamma_u * \tilde{c}^{(t)} + (\Gamma_f) * c^{(t-1)}$$

$$\Gamma_o * c^{(t)} = a^{(t)}$$

**SİYAH BEYAZ KÖPEK
BAR ÜZERİNDEN
ATLIYOR**



**PEMBE KIYAFETLİ KIZ
HAVADA ZIPLIYOR**



**TURUNCU GÜVENLİK YELEĞİ
GİYEN YAPI İŞÇİSİ YOLDA
ÇALIŞIYOR**



**MAVİ DALIŞ KIYAFETLİ
ADAM DALGA ÜSTÜNDE
SÖRF YAPIYOR**

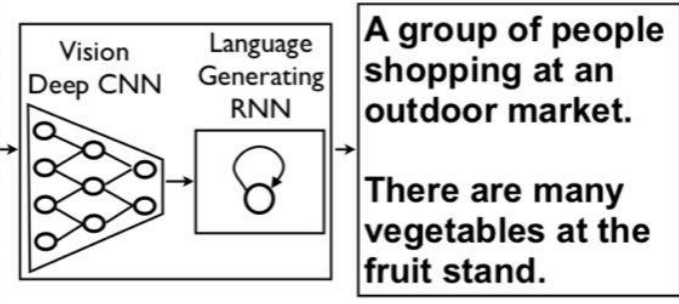


**İKİ GENÇ KIZ LEGO
OYUNCAĞI İLE
OYNUYOR**



**SİYAH TİŞÖRT GİYEN
ADAM GİTAR ÇALIYOR**



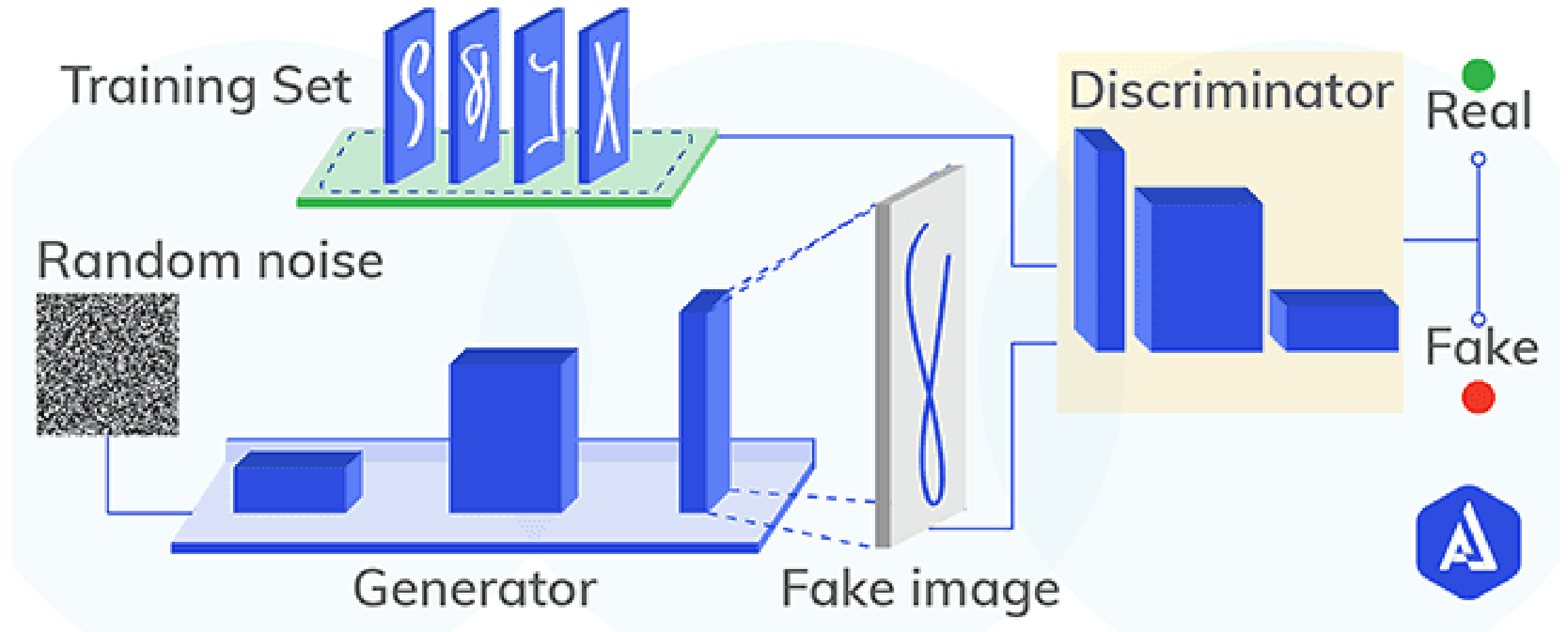


M. Kuyu, A. Erdem ve E. Erdem.
Altsözcük Öğeleri ile Türkçe Görüntü Altyazılama.
SİU 2018

Yarış pistinde virajı almakta olan bir yarış arabası

ÜRETİCİ ÇEKİŞMELİ AĞLAR





YAPAY ZEKA'NIN 3 SENELİK GELİŞİMİ



2014



2015



2016



2017



STİL TRANSFERİ

Monet ↔ Photos



Monet → photo

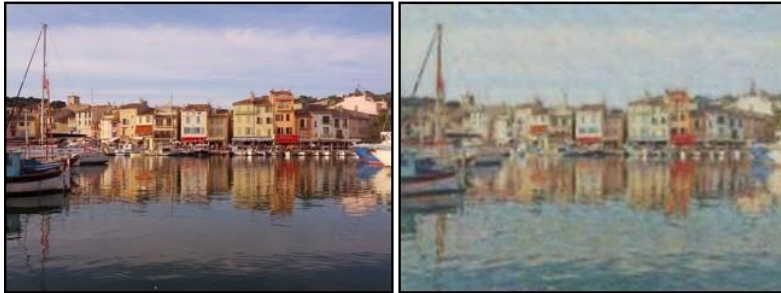


photo → Monet

Zebras ↔ Horses



zebra → horse



horse → zebra

Summer ↔ Winter



summer → winter



winter → summer



Photograph



Monet



Van Gogh



Cezanne



Ukiyo-e

Derin Öğrenme ile Artistik Stil Transferi

You highlighted

🎨 Çektiğiniz fotoğrafın ya da çizdiğiniz bir resmi tüm zamanların sizce * en başarılı ressamı resmetseydi neye benzerdi, merak etmez misiniz?



Ayyüce Kızrak

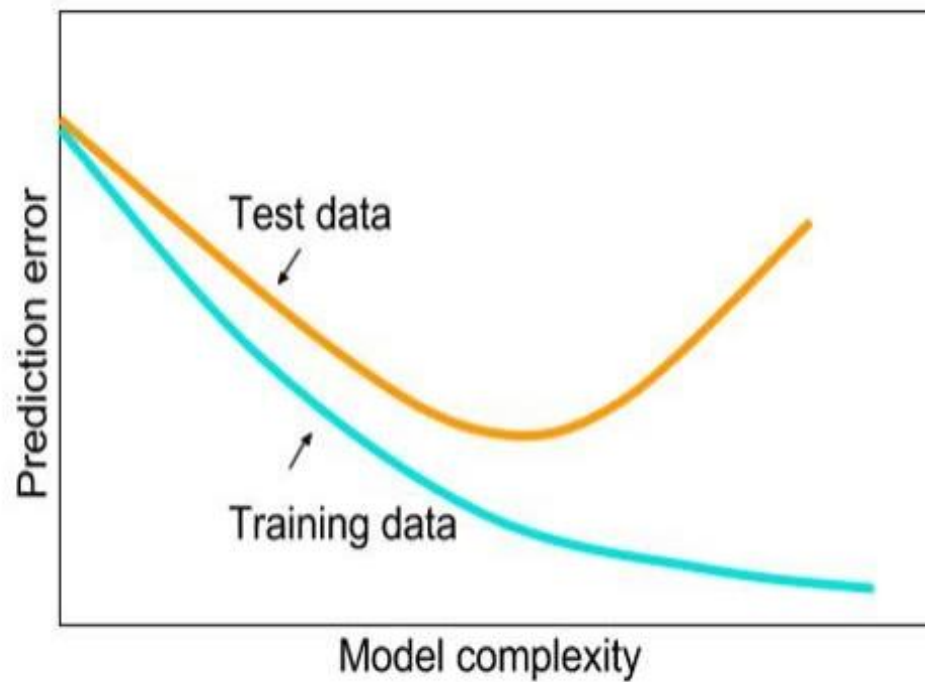
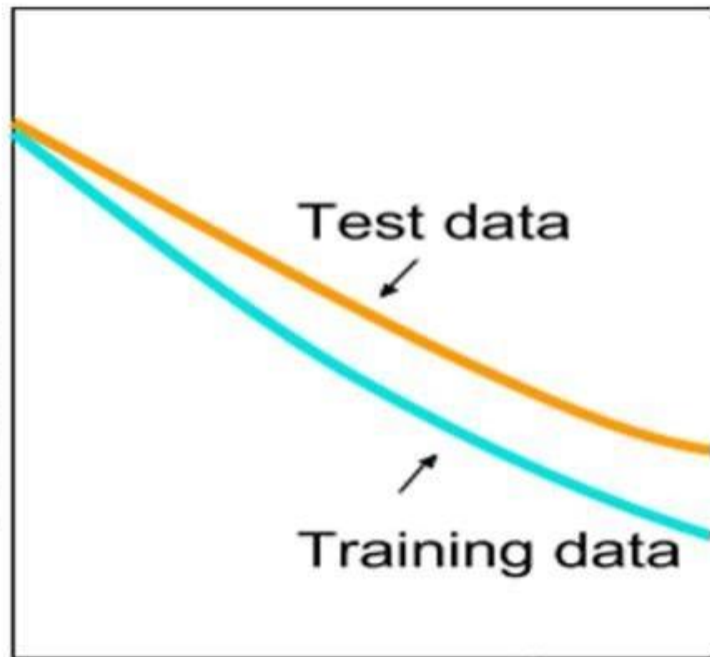
Nov 26, 2018 - 7 min read

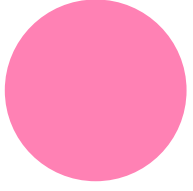
Peki bu, sanatçıyı kopyalamak mıdır, ölümsüzleştirmek midir?



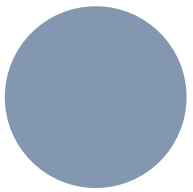
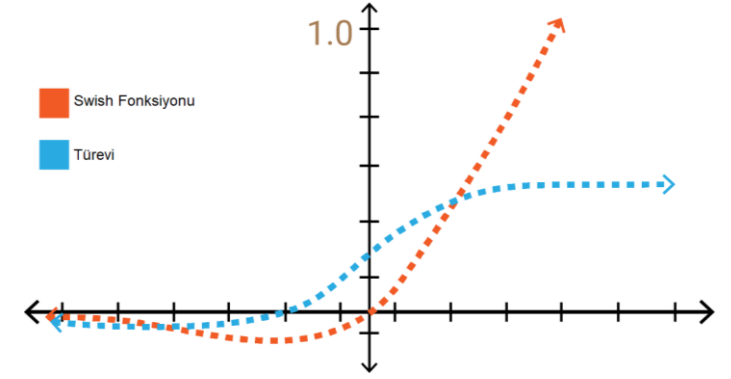
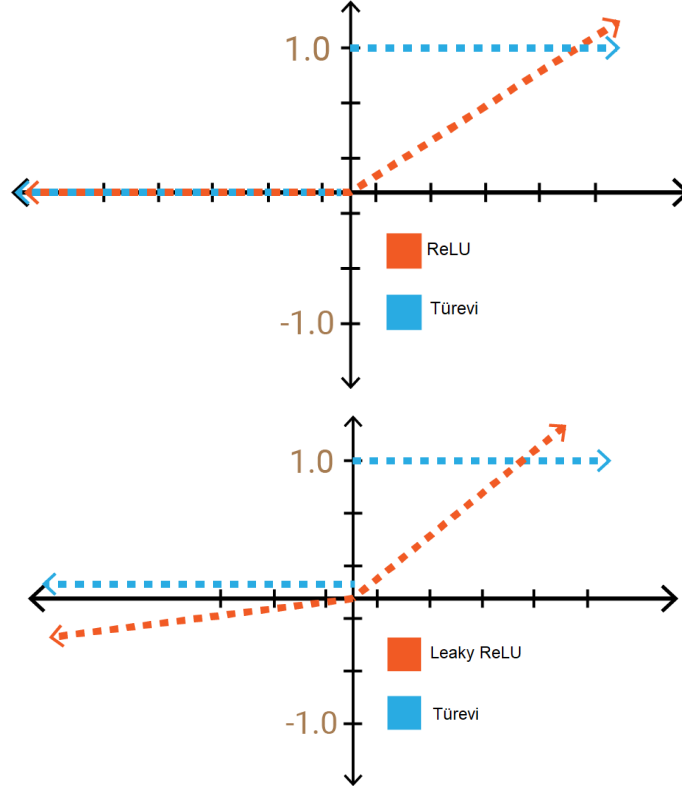
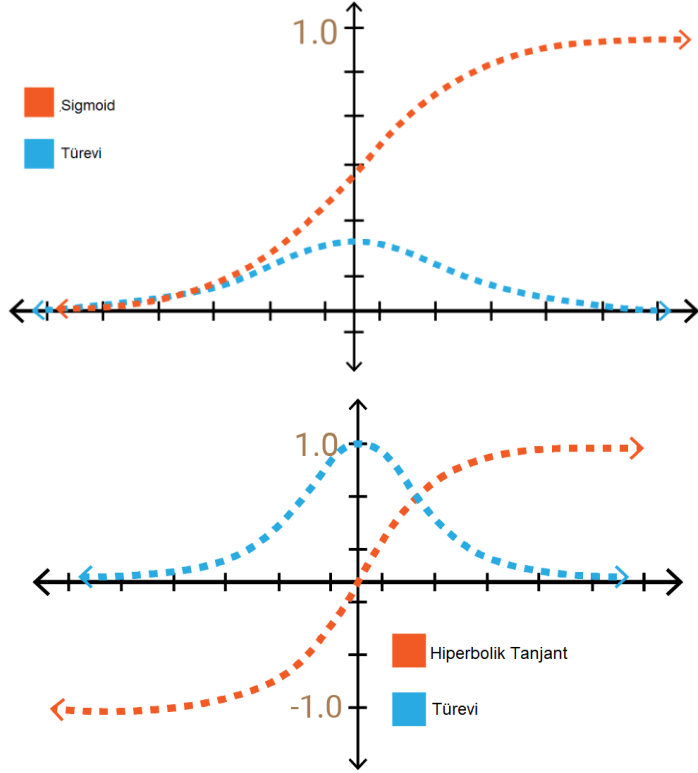
DERİN ÖĞRENME VE KARMAŞIK SİSTEMLER

Prediction error



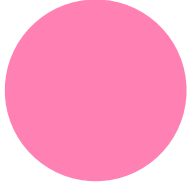


Aktivasyon Fonksiyonunun Doğru Seçilmesi

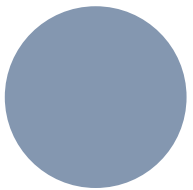
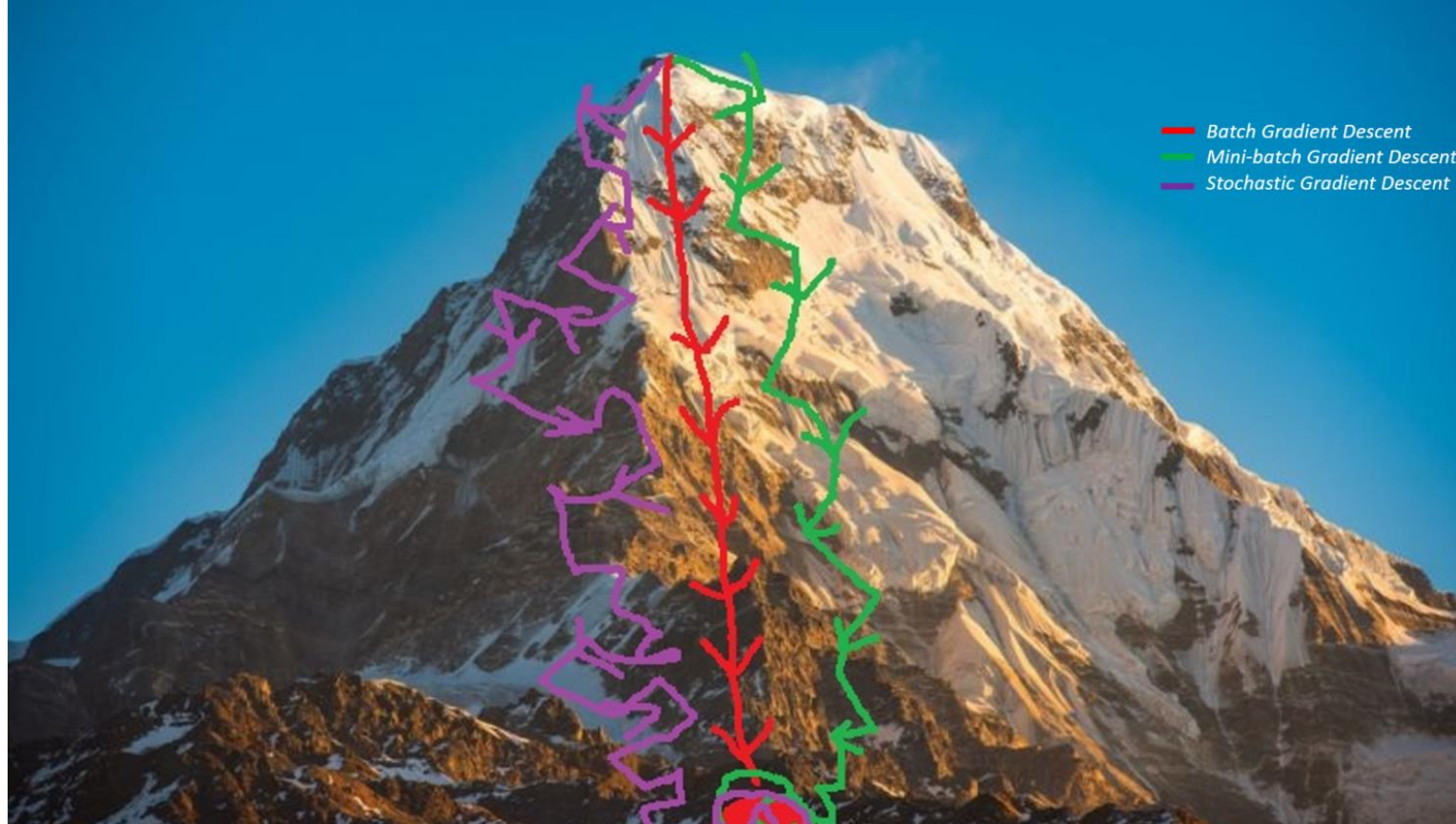


ÇÖZÜM: İŞLEM YÜKÜNÜ ÇOK ARTIRMADAN EN İYİ ÖĞRENEN FONKSİYONU BELİRLEMEK

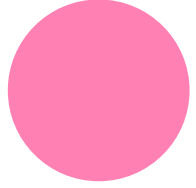
Karmaşık Sistemler ve Veri Bilimi Yaz Okulu 2019-M. Ayyüce Kızrak



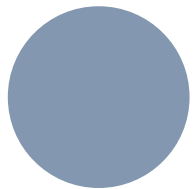
Optimizasyon Yönteminin Seçilmesi



ÇÖZÜM: ÖDÜNLEŞİM

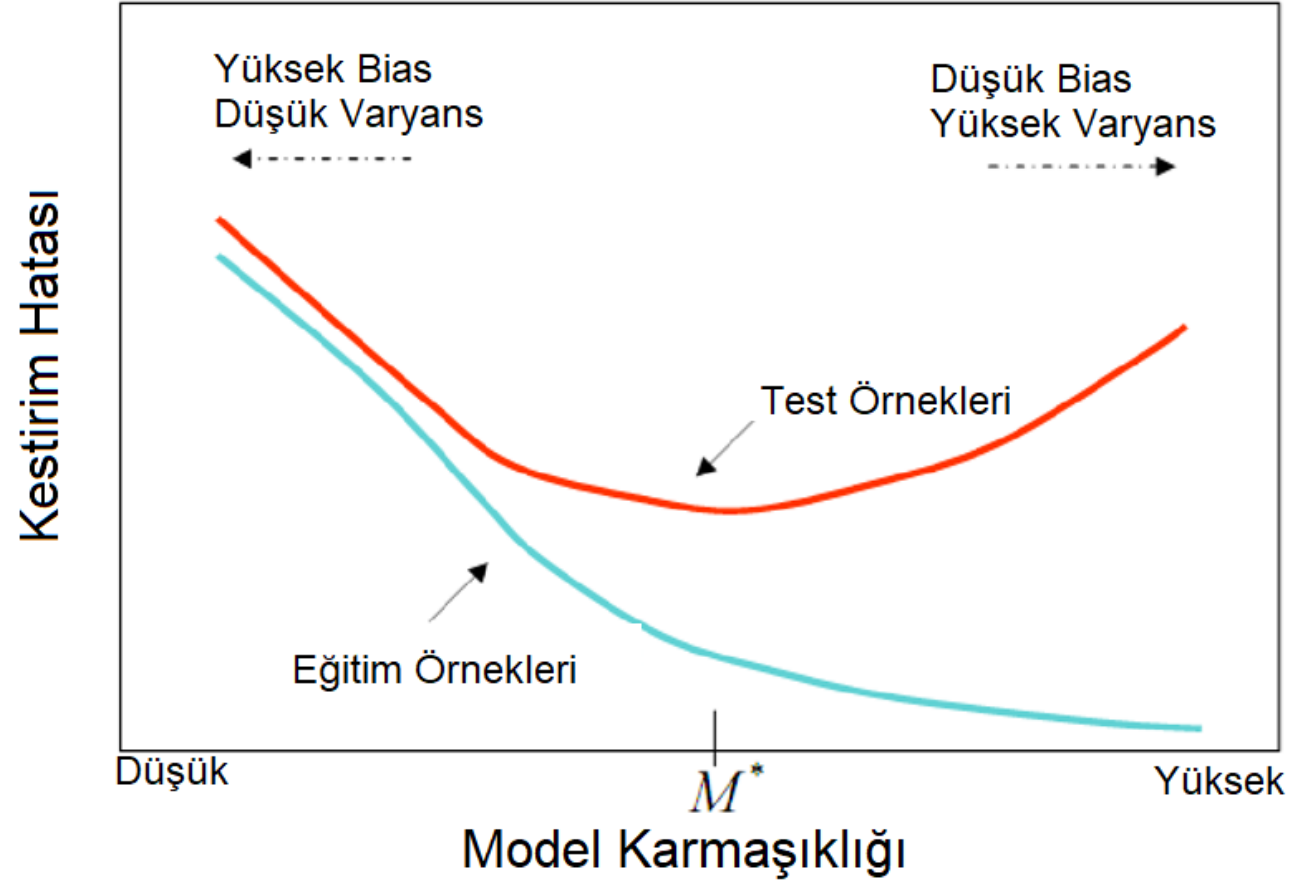


Bias vs. Variance

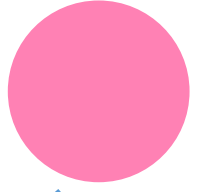


ÇÖZÜM: GENELLEŞTİRİLEBİLİR, AÇIKLANABİLİR VE BAŞARILI BİR MODEL

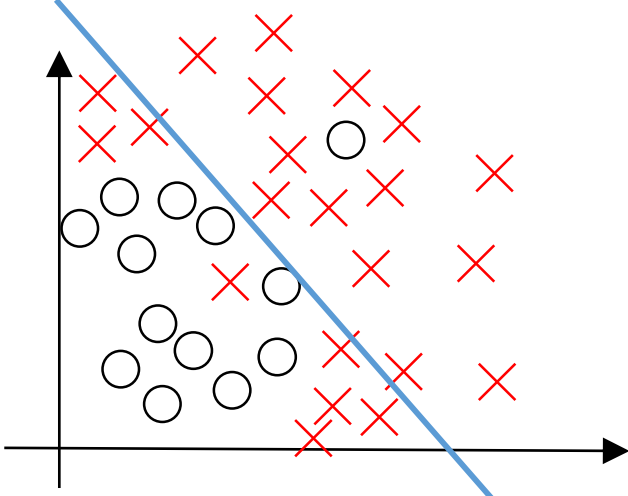
Bias vs. Variance



ÇÖZÜM: GENELLEŞTİRİLEBİLİR, AÇIKLANABİLİR VE BAŞARILI BİR MODEL

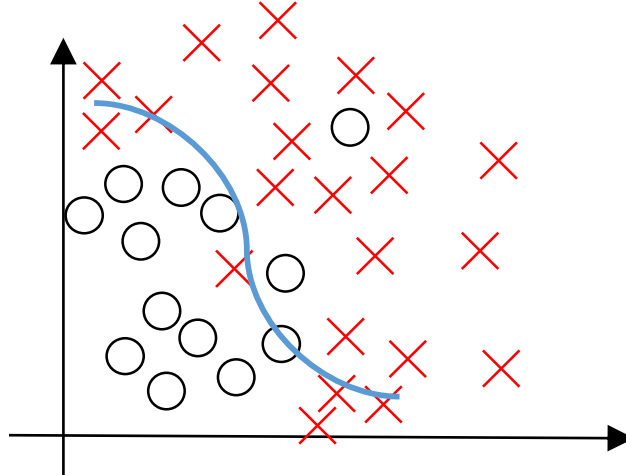


Bias vs. Variance

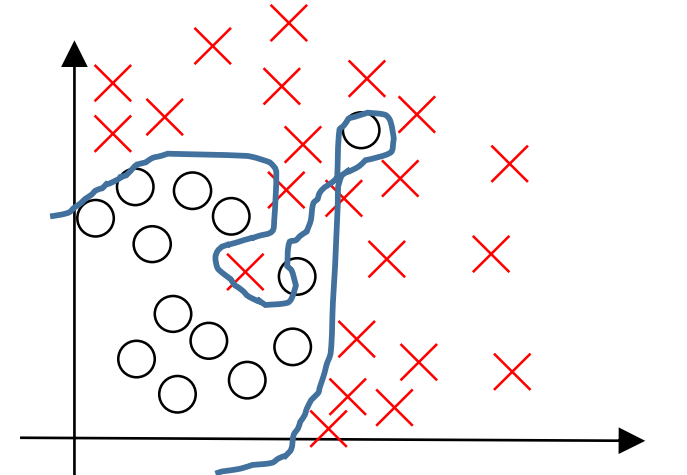


yüksek bias

Az Uydurma



“tam kararında”

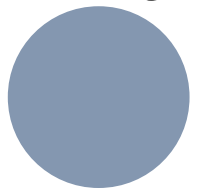


yüksek varyans

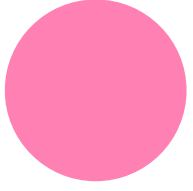
Aşırı Uydurma

Bias = modelin tahminlerindeki sistematik hata.

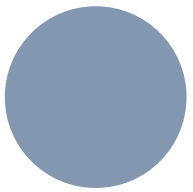
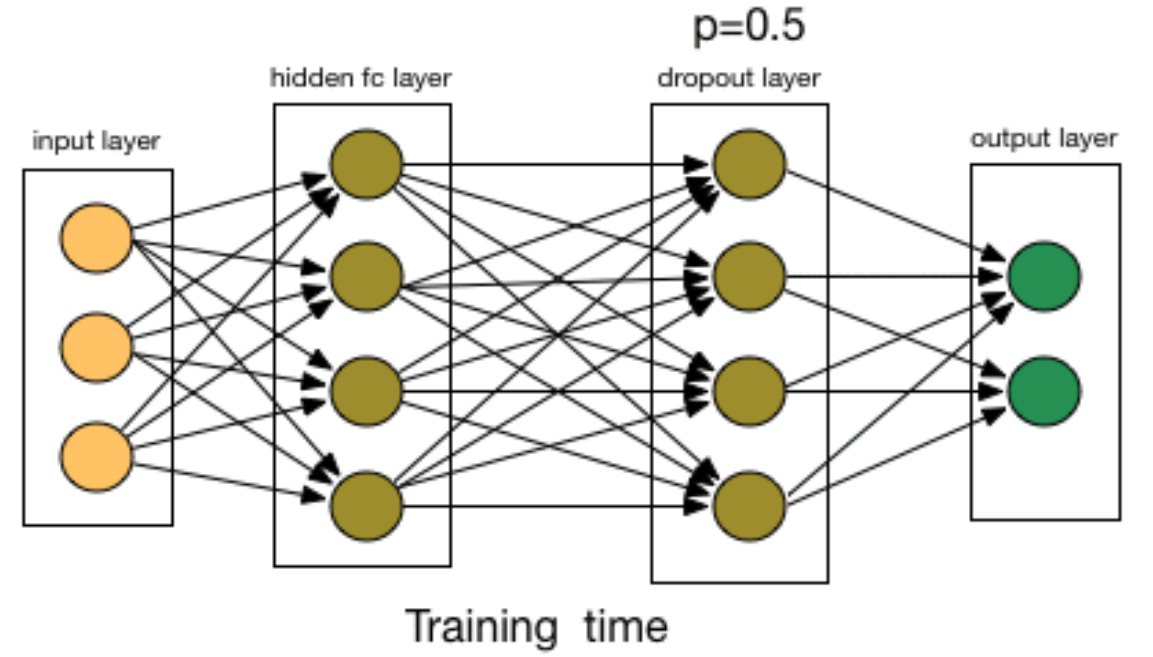
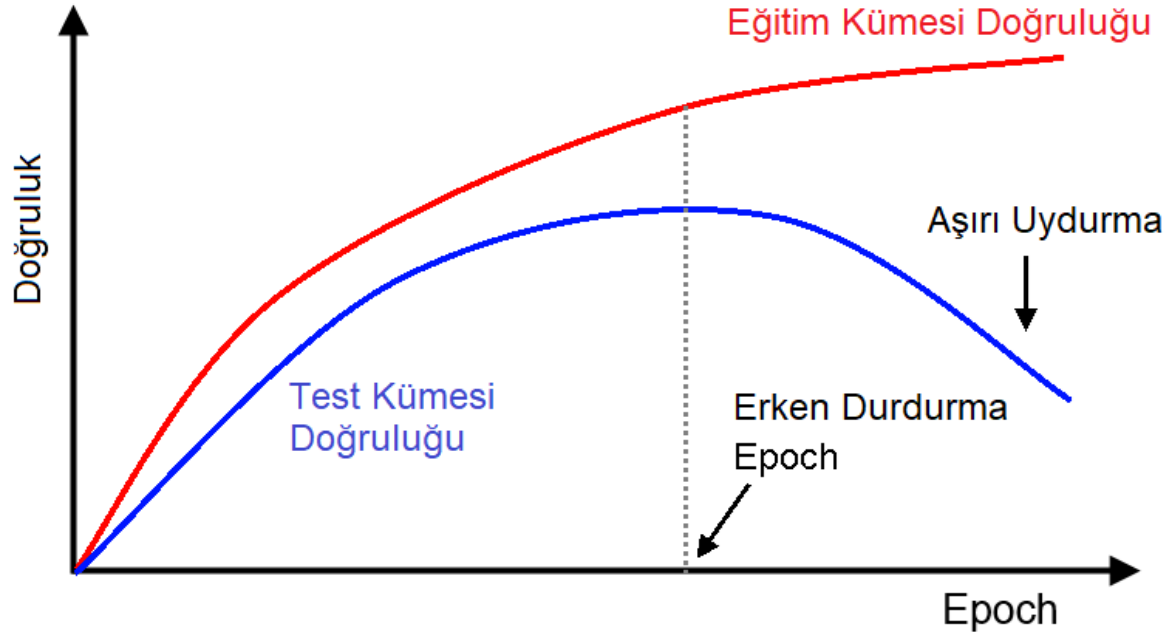
Variance = eğitim setinde örnekleme gürültüsü tahminlerdeki gürültüye ne neden olur.



ÇÖZÜM: GENELLEŞTİRİLEBİLİR, AÇIKLANABİLİR VE BAŞARILI BİR MODEL



Aşırı uydurma- Az uydurma (Overfitting vs. Underfitting)



ÇÖZÜM: SEYRELTME-ERKEN DURDURMA-GÜRÜLTÜ EKLEME

Yapay Öğrenme Modeli Geliştirme Püf Noktaları El Özeti

Yapay Öğrenme Modeli Geliştirirken Genellikle Karşılaşılan Problemler ve Olası Çözümleri

Developer Summit 2019 organizasyonunda yapmış olduğum sunumu öylece paylaşmak yerine kaynaklarla destekleyerek bir özet sunmak istedim. Hem katılamayanlar için de belki "hap bilgi" dediğimiz bir yönlendirme olur diye düşündüm. Elbette ki her bir başlık kendi başına uzun uzun anlatılabilir ve zengin içeriğe sahiptir. Bu kaynak, ihtiyacı olanları hızlıca yönlendirmek için hazırlanmıştır. Verdiğim bağlantıları takip ederek uygulamalı deneylerine ve daha açıklamalı kaynaklara da ulaşabilirsiniz!

Tek sayfa *Yapay Öğrenme Modeli Geliştirme Püf Noktaları El Özeti*'ni indirmek için [PDF](#) veya [PNG](#) tıklayınız!

[Aktivasyon Fonksiyonları için detaylı açıklama ve Google Colab üzerinde uygulama örneği](#)

[Aktivasyon Fonksiyonları için videolu anlatım](#)

[Optimizasyon Algoritmalarının Uygulamalı Karşılaştırılması](#)

[Transfer Öğrenme için detaylı açıklama ve uygulama örnekleri](#)

GENEL KÜLTÜR OKUMA ÖNERİLERİ



Okuma Önerileri Reading List



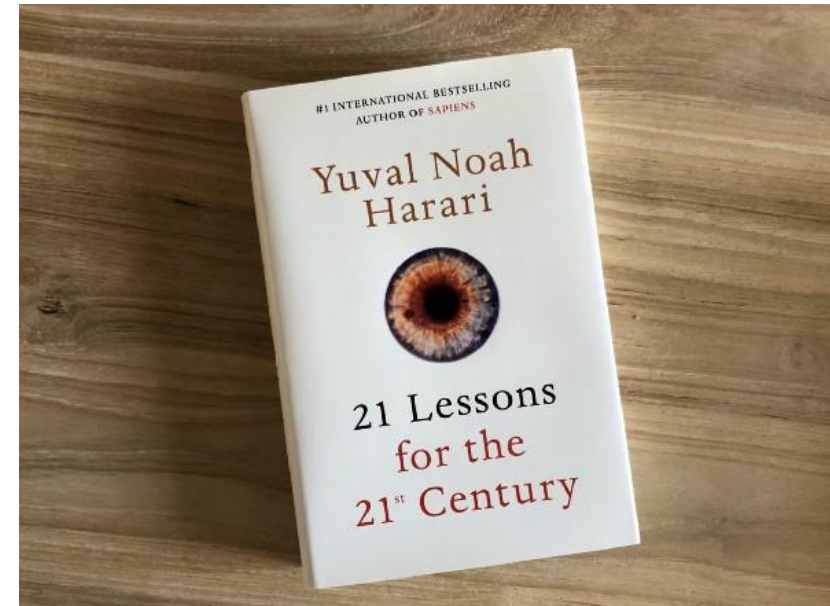
Ayyüce Kızrak

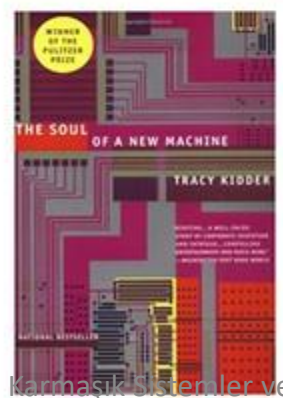
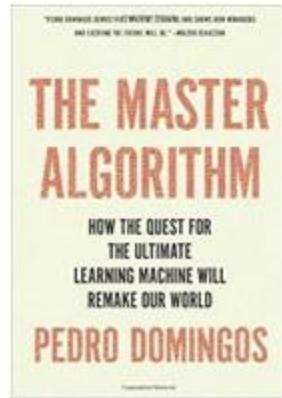
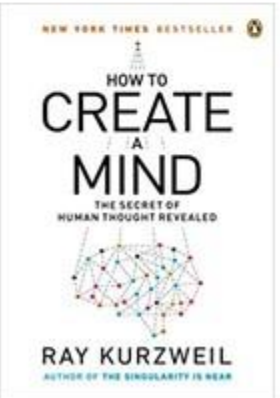
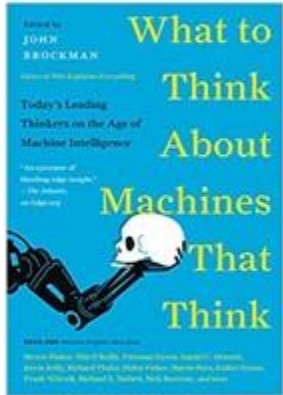
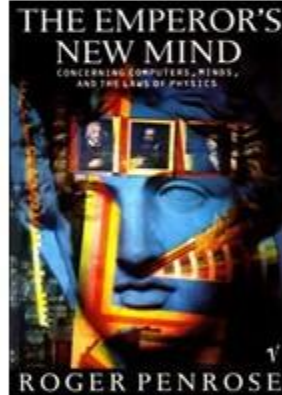
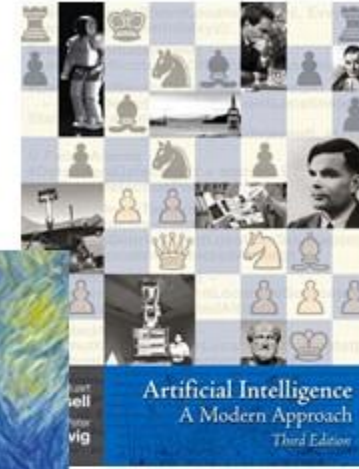
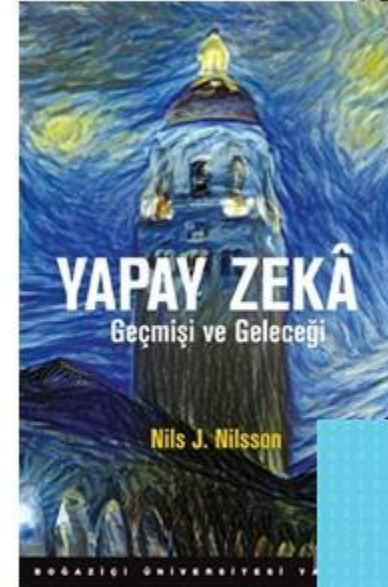
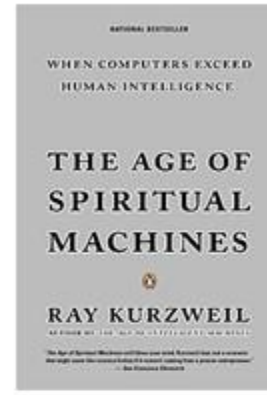
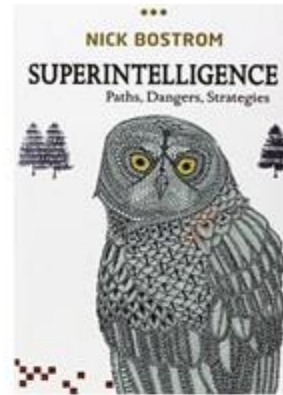
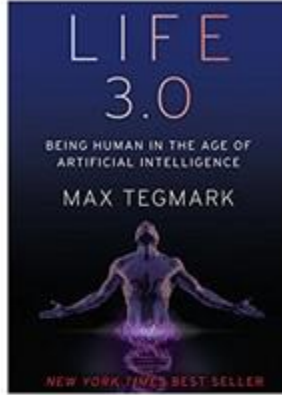
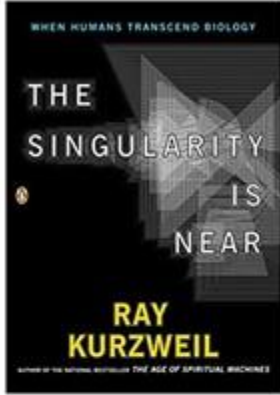
Mar 31 · 7 min read

"Bilimi açıklamamak bana ahlaksızlık gibi geliyor, aşık olunca bunu tüm dünyaya duyurmak istersiniz."

"Not explaining science seems to me perverse. When you're in love, you want to tell the world."

~Carl Sagan





Karmaşık Sistemler ve Veri Bilimi
Yaz Okulu

TEŞEKKÜRLER



www.ayyucekizrak.com

M. AYYÜCE KIZRAK