

RandomTalksEducation
- Seminar Series -



Yapay Zeka

“Tarihi ve Güncel Durum”

Yapay Sinir Ağlarının Sıfırdan Keşfi

Prof. Dr. Bekir Taner Dinçer

Muğla Sıtkı Koçman Üniversitesi
Mühendislik Fakültesi
Bilgisayar Mühendisliği Bölümü



1

Anlatım Akışı

1. Bölüm: Yapay Zeka alanında güncel durum

Geçmişten günümüze gelişmeler:

- Güncel durum
- Tarih tekerrür edecek mi?

2. Bölüm: Yapay Sinir Ağlarının Sıfırdan Keşfi

Çokça sorulan sorulara cevaplar:

- 1.Sınava 12 saat çalışırsam kaç puan alırım?
- 2.Kaç saat çalışırsam dersten geçerim?
- 3.Hangi kaynaklardan sınava çalışayım?



2

Ben Kimim?



3

Eğitim

İSTATİSTİK

Lisans

BİLGİSAYAR BİLİMLERİ

Lisans Üstü



4

2

İlgili Biyografi: Eğitim



5

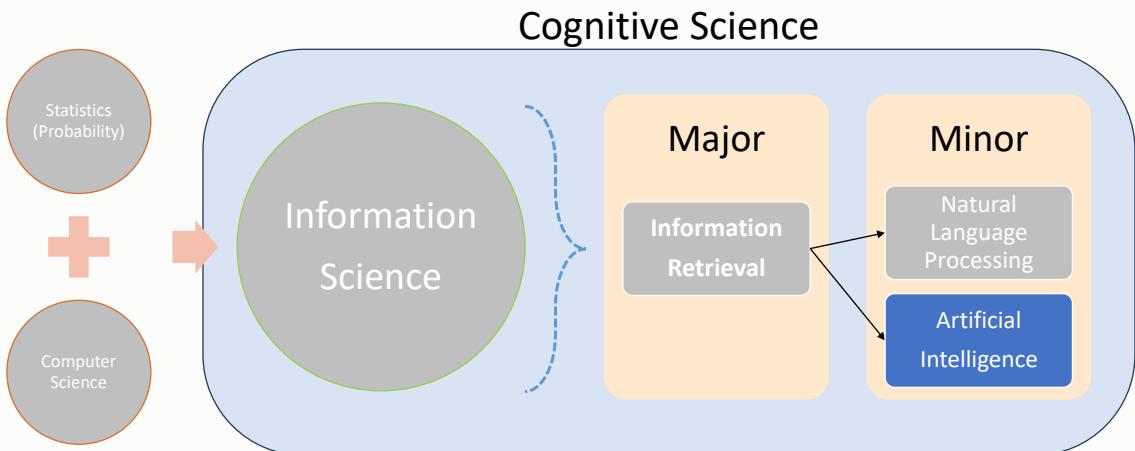
İlgili Biyografi: Eğitim

Yapay
Zeka



6

Uzmanlık Alanları



7

YZ Alanının Güncel Durumu

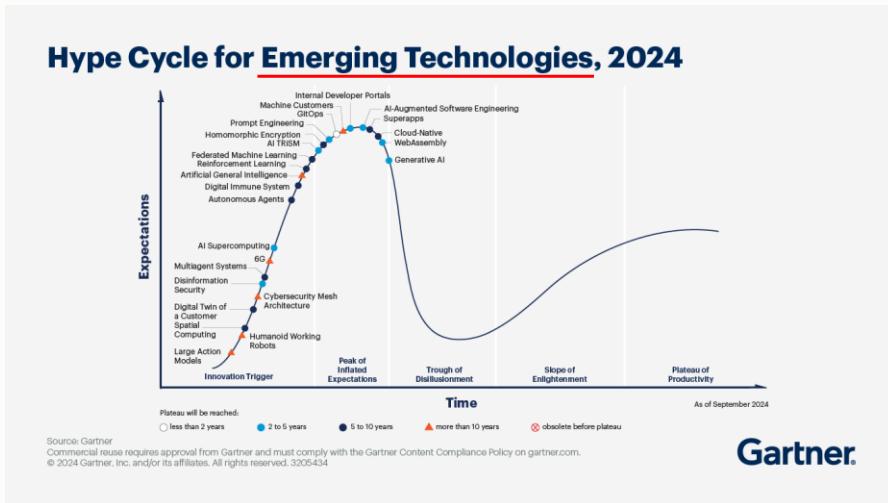
Teknolojilerin Olgunluk Seviyeleri



8

2024 Yılı Beklenti Raporu

<https://www.gartner.com/en>



9

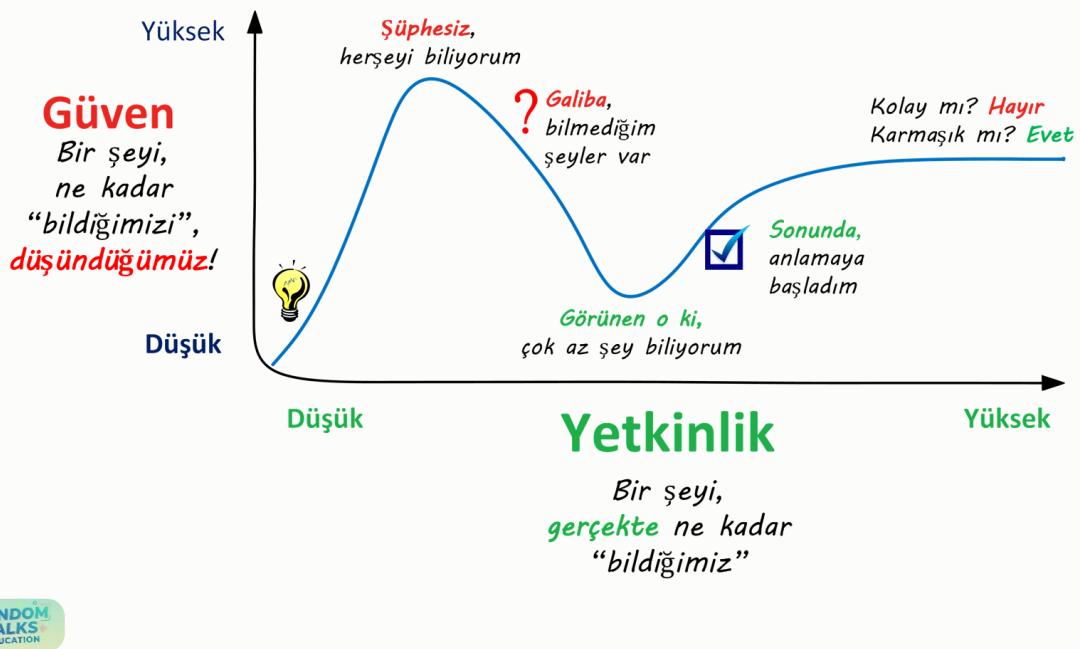
Dunning-Kruger Etkisi

2000 Yılı Nobel Ödülü



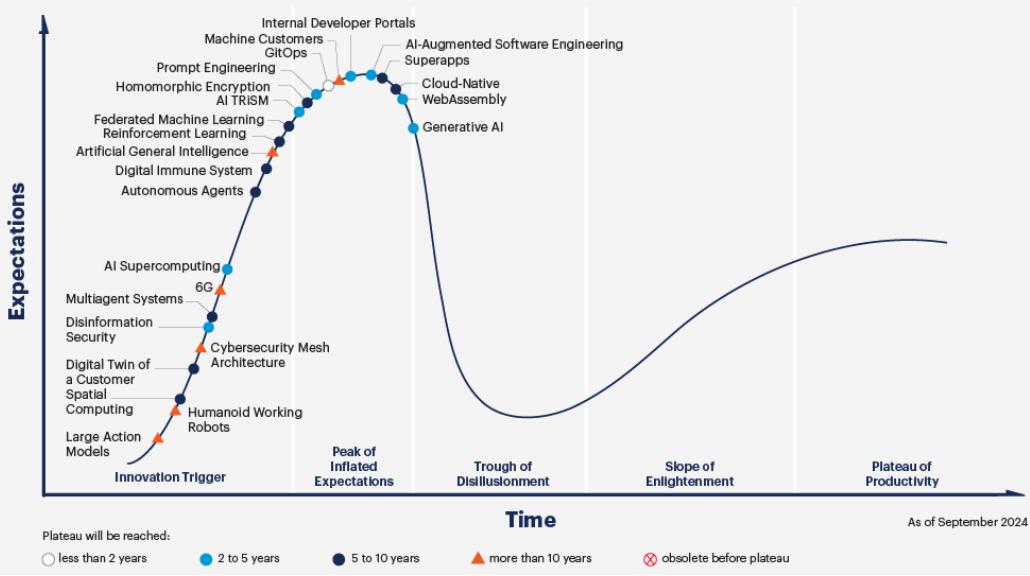
10





11

Hype Cycle for Emerging Technologies, 2024

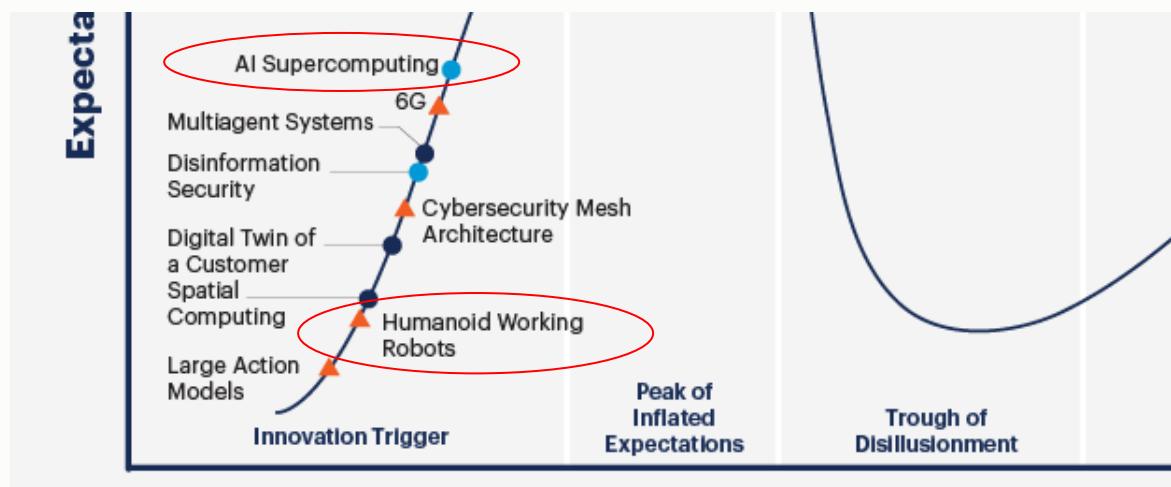


12



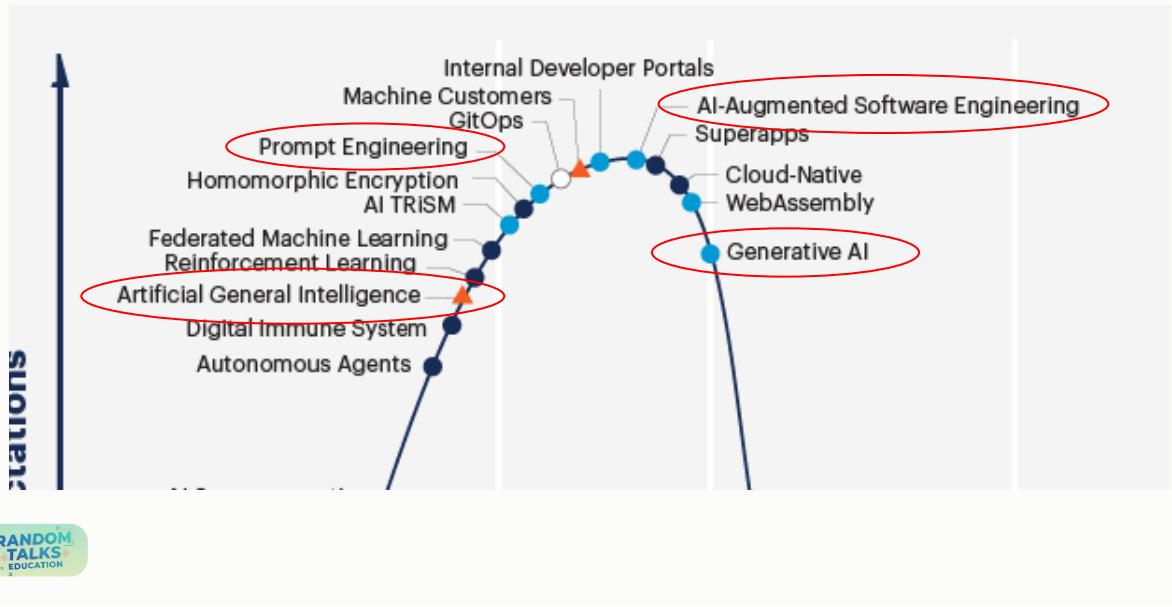
13

Yükselenler



14

Tepe Noktasındakiler



15

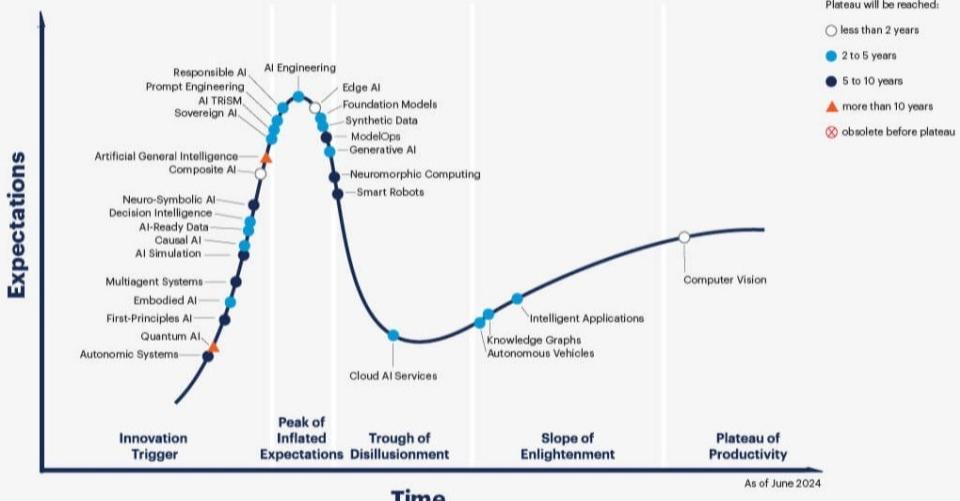
Yapay Zeka Alanı

Güncel Durum



16

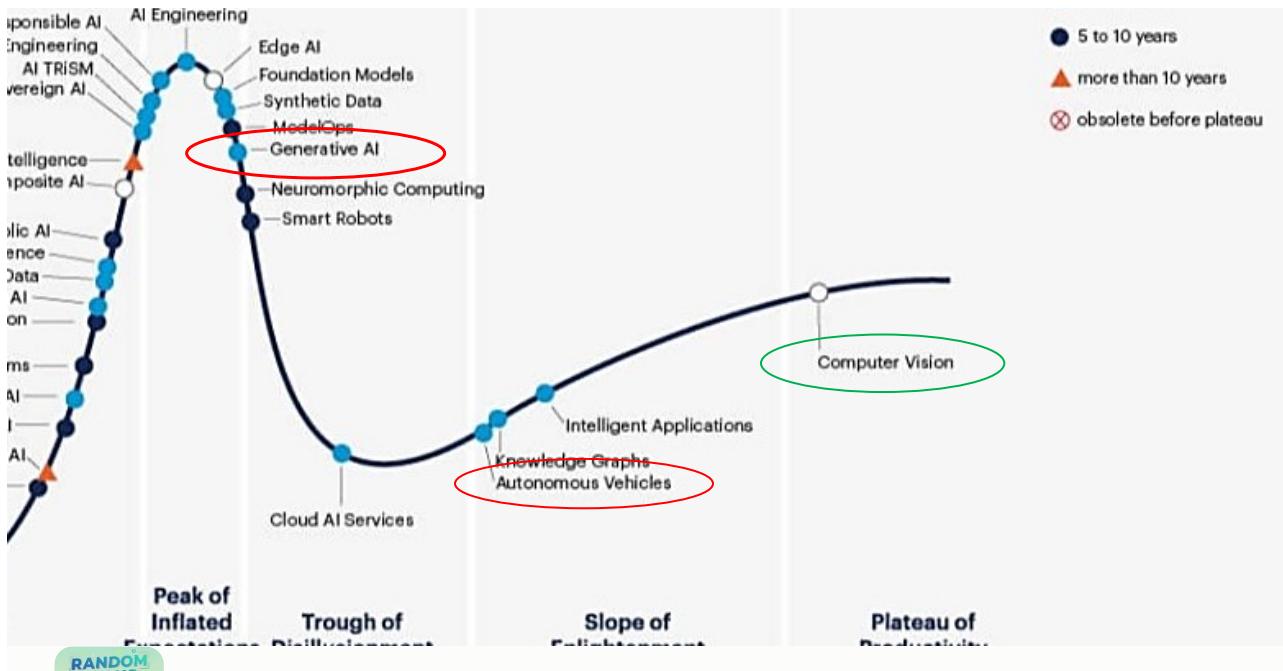
Hype Cycle for Artificial Intelligence, 2024



Source: Gartner
 Commercial reuse requires approval from Gartner and must comply with the
 Gartner Content Compliance Policy on gartner.com.
 © 2024 Gartner, Inc. and/or its affiliate(s). All rights reserved. GTS_3282450

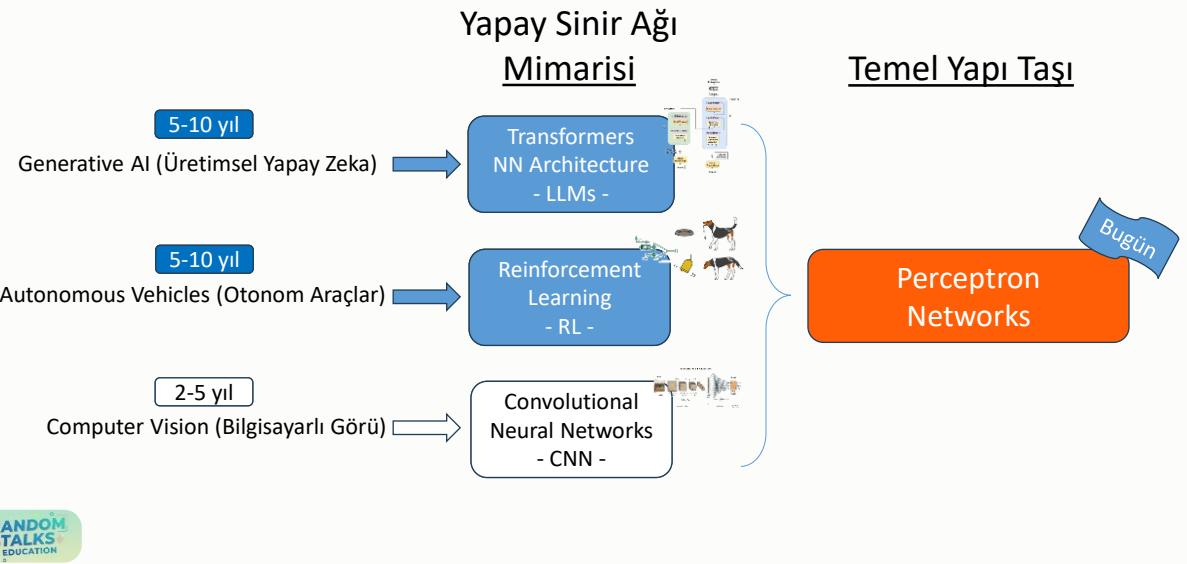
Gartner®

17



18

Teknolojiler ve Yapay Sinir Ağlı Mimarileri

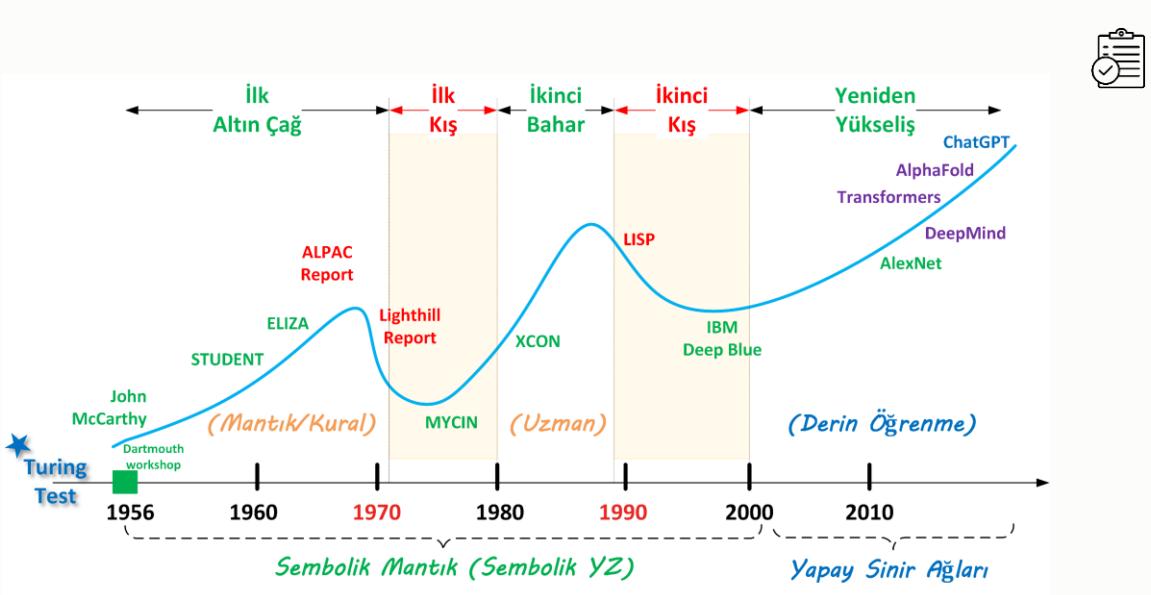


19

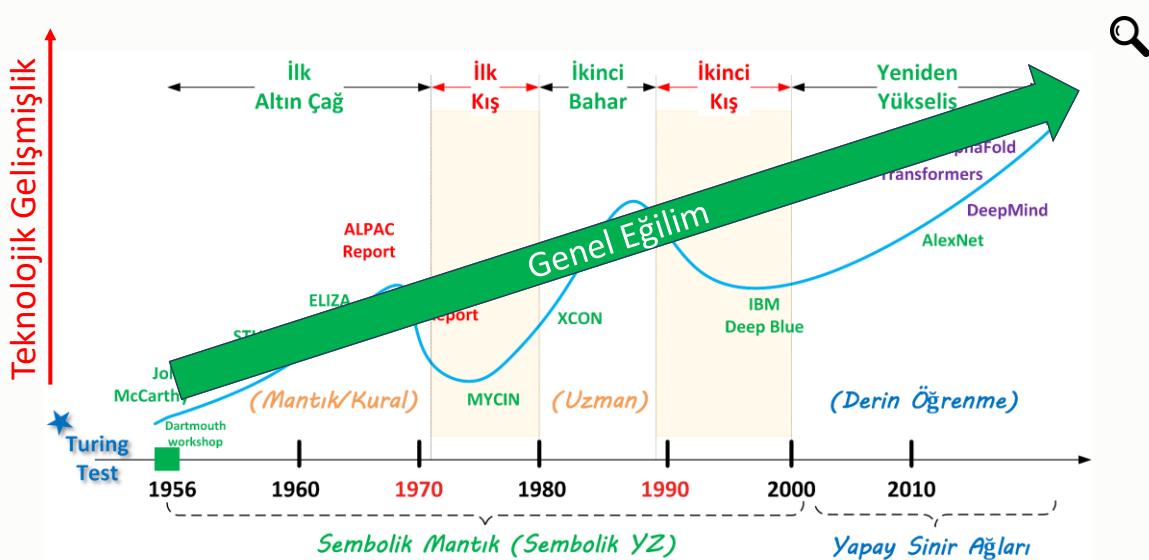
Tarih tekerrür edecek mi?

Güncel Durum

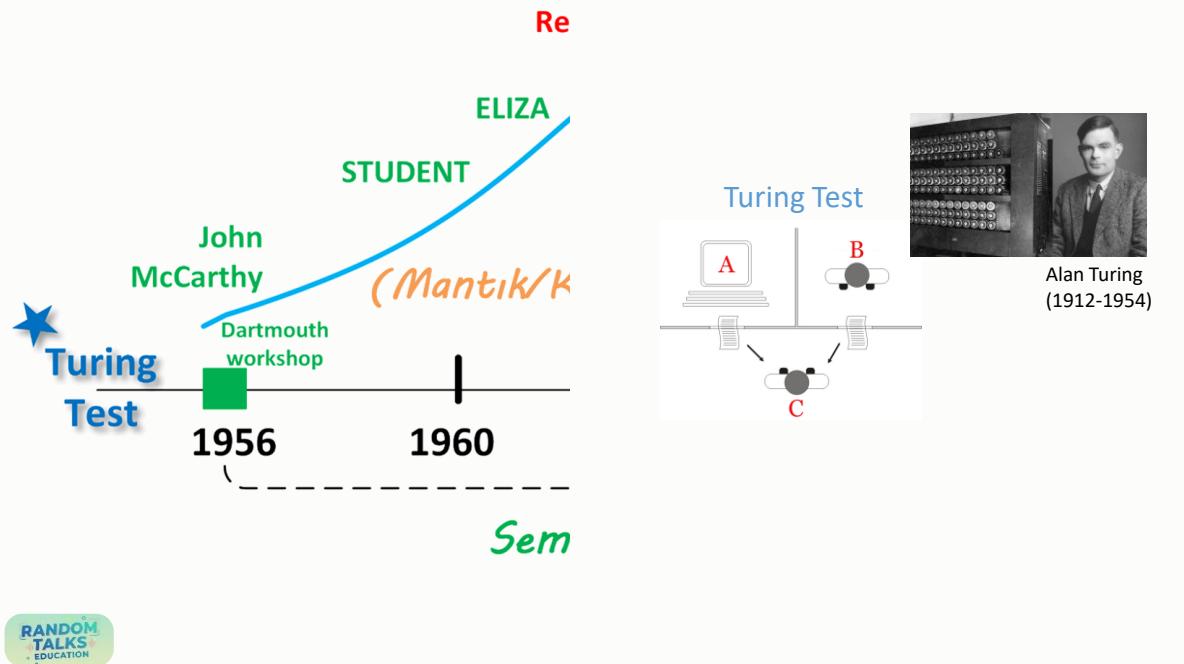
20



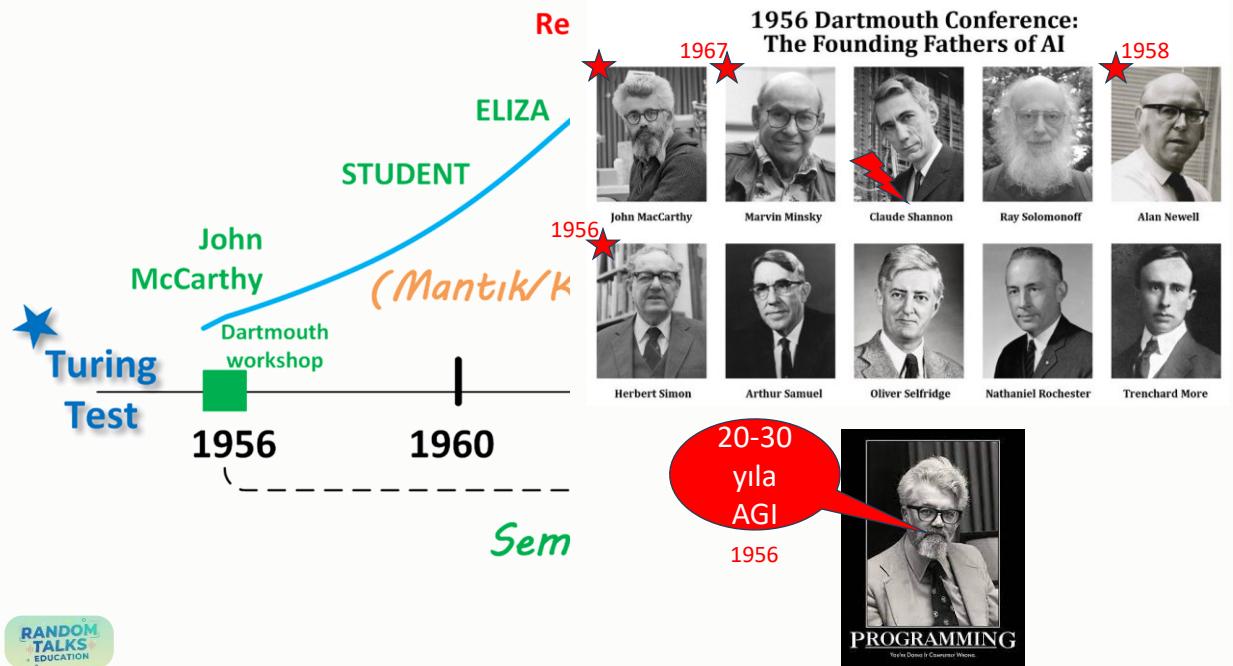
21



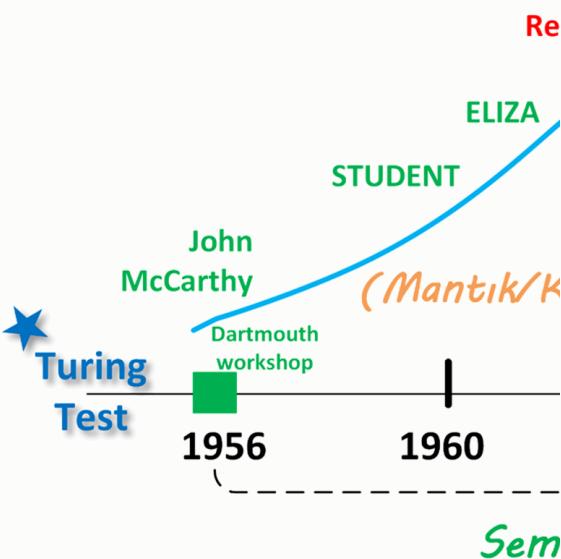
22



23



24



Daniel Bobrow

STUDENT

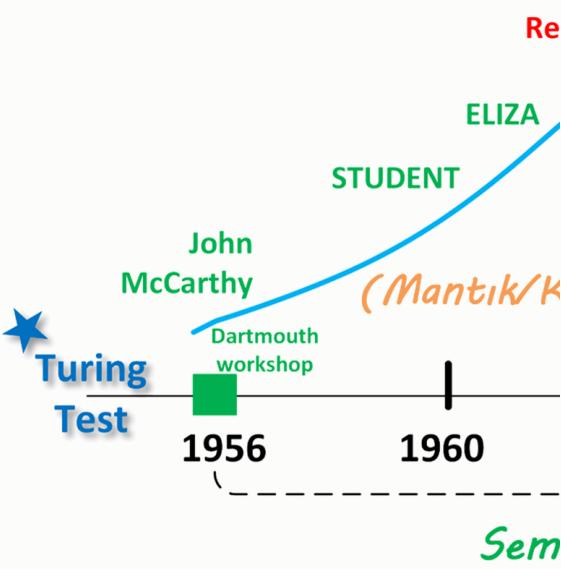
STatistical UNDERstander

PhD thesis,

"Natural Language Input for a Computer Problem Solving System," in 1964, MIT.

Rol	Girdi/Cıktı
Kullanıcı Girdisi	"Müşteri sayısı, reklam sayısının iki katıdır. Müşteri sayısı 10'dur. Reklam sayısı kaçtır?"
STUDENT Çıktısı (Dahili Çeviri)	Denklem 1: MÜŞTERİ SAYISI = 2 * REKLAM SAYISI Denklem 2: MÜŞTERİ SAYISI = 10 Soru: REKLAM SAYISI = ?
STUDENT Çıktısı (Çözüm)	Reklam sayısı 5'tir.

25



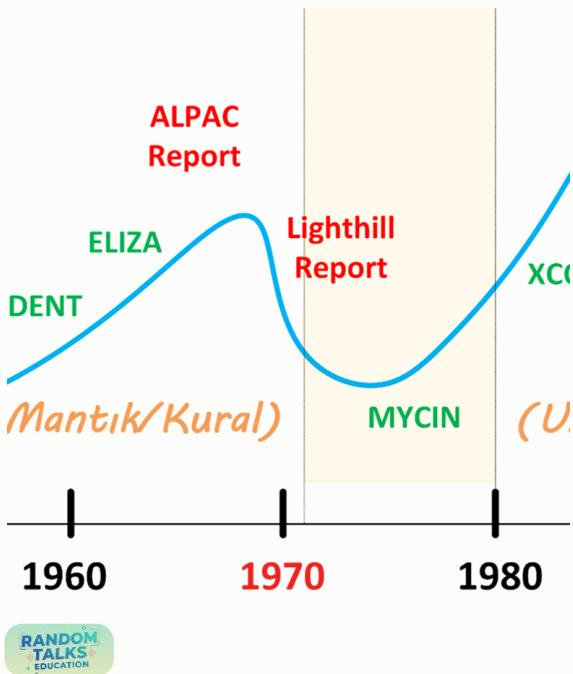
ELIZA

(1966)

Joseph Weizenbaum
(1923-2008)

Konuşmacı	Diyalog	ELIZA'nın Stratejisi (Çeviri)
Kullanıcı	Son zamanlarda kendimi çok kötü hissediyorum. Sanırım işim çok zorlu.	Anahtar Kelime: "Hissediyorum" → Duygular hakkında bir soruya yanıt verir.
ELIZA	Bana o hislerinizden daha çok bahseder misiniz?	Anahtar Kelime: "aile" → Aile ilişkisine odaklanır.
Kullanıcı	Aileme artık yeterince zaman ayıramıyorum ve bu beni üzüyor.	Anahtar Kelime: "anenm" → Odağı belirli bir aile üyesine kaydırır.
ELIZA	Ailenize yeterince zaman ayıramamanız siz üzüyor mu?	Yansıtma/Yeniden İfade Etme → Kullanıcının ifadesini soru olarak tekrarlar.
Kullanıcı	Evet, kesinlikle. Annem kariyerime çok odaklandığım için beni hep eleştirdi.	Anahtar Kelime: "Ben" → Kullanıcının kendisiyle ilgili ifadesini daha fazla araştırmak için bir soruya dönüştürür.
ELIZA	Bana annenizden daha çok bahseder misiniz?	
Kullanıcı	Hep daha erken evde olabileceğim bir iş bulmam gerektiğini söyledi.	

26



1966: ALPAC Raporu Darbesi

(Automatic Language Processing Advisory Committee)

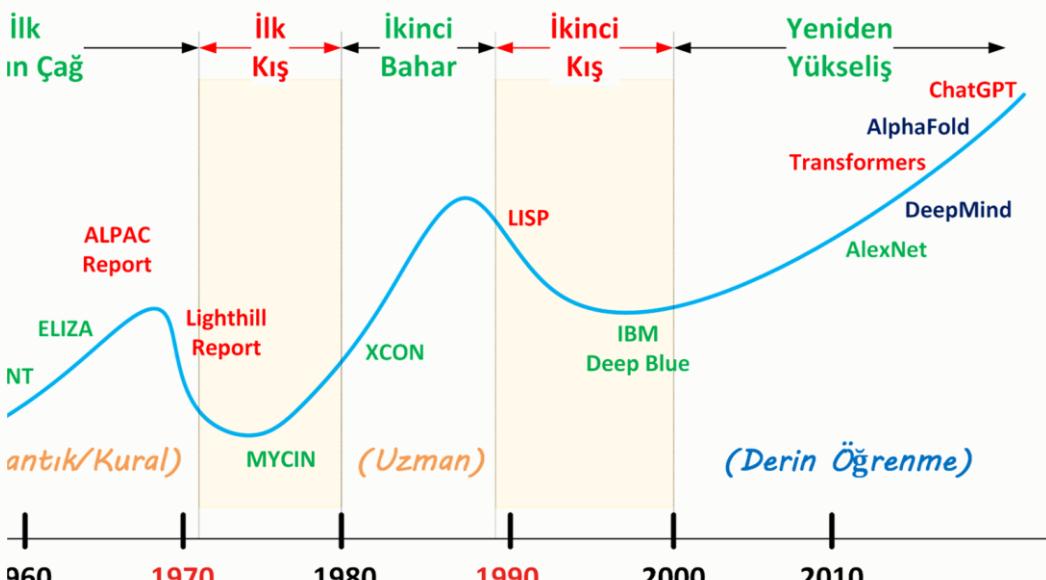
- ALPAC raporu, Makine Çevirisi alanındaki ilerleme eksikliğini eleştirdi ve makine çevirisinin insan çevirisinden daha pahalı, daha az doğru ve daha yavaş olduğu sonucuna vardı. Bu, ABD hükümetinin Doğal Dil İşleme (NLP) fonlarında erken bir kesintiye yol açtı.

1973: Lighthill Raporu (Kışın Köşe Taşı)

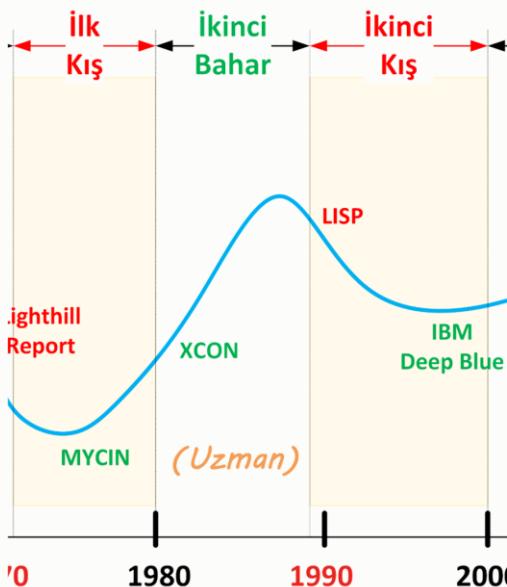
- İngiltere hükümeti tarafından görevlendirilen Lighthill Komisyon Raporu, yapay zeka araştırmalarını "iddialı hedeflerine" ulaşmadaki "mutlak başarısızlığı" nedeniyle sert bir şekilde eleştirdi ve temel değerini sorguladı.

→ Fonlar Durduruldu

27



28

**MYCIN,**

- bir Uzman Sistemdir (Expert System) ve tıp alanındaki ilk başarılı yapay zeka uygulamalarından biridir.



Doktora Tezi (1972)
Stanford Üniversitesi, Kaliforniya.

Edward Shortliffe

XCON (eXpert CONfigurer)

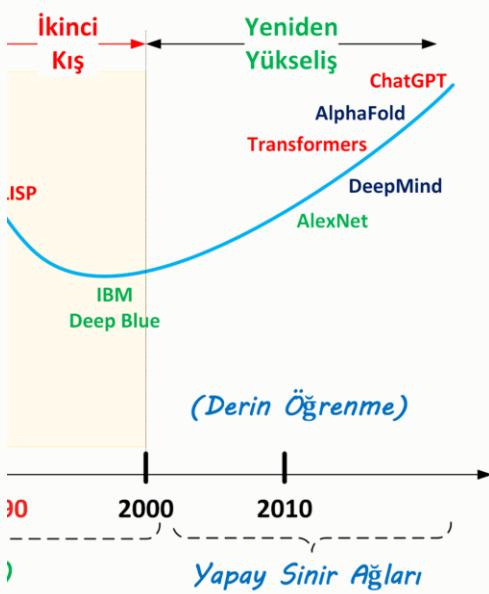
- DEC'in ürettiği büyük VAX bilgisayar sistemleri için müsteri siparişlerini otomatik olarak doğru bileşenlerle yapılandırmaktı.

Carnegie Mellon Üniversitesi, 1978
DEC Inc., 1980



- Kanıtlanmış ROI (Yatırım Getirişi)
 - yapay zeka şirketlerinin kurulmasına,
 - LISP makinesi endüstrisinin yükselişine ve
 - yapay zeka araştırmalarına yeniden yatırımlar
- DEC'in ürün yelpazesi sürekli değiştiği için,
 - XCON'un kural tabanı da sürekli olarak güncellenmek zorundaydı ve
 - bu, zamanla yüksek bakım maliyetlerine ve kırılganlığa yol açtı.

29



IBM Deep Blue
(1997)

→ Sembolik YZ Çalışmalarının Sonu

- AlexNet (2012)** – Görüntü İşleme – MLPs (-> CNN)

Alex Krizhevsky, (Ilya Sutskever, Geoffrey Hinton)

→ (Derin) Yapay Sinir Ağlarının Yükselişi

- (Google) DeepMind (2012)** – Atari Oyunları, (Derin) RL
- (Google) Transformers (2017)** – Doğal Dil İşleme - LLMs
- (Google) AlfaFold (2017)** – Protein Katlama - DeepMind

30

“Winter is coming!”

AGI Vaatlerinin Yerine Getirilememesi:

Büyük Dil Modellerinin (LLM'ler) Yapay Genel Zeka'nın (AGI) eşiğinde olduğu yönündeki aşırı abartılı söylemler.



!!! Google, Apple, Meta, Microsoft !!!

!!! Akademi !!!

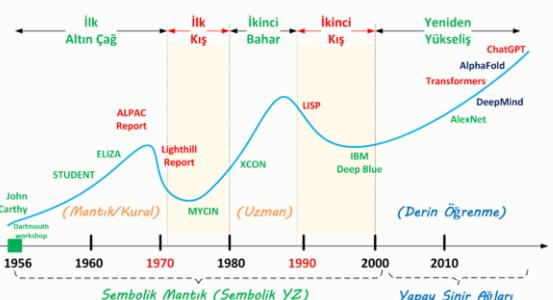


31

“Winter is coming!”

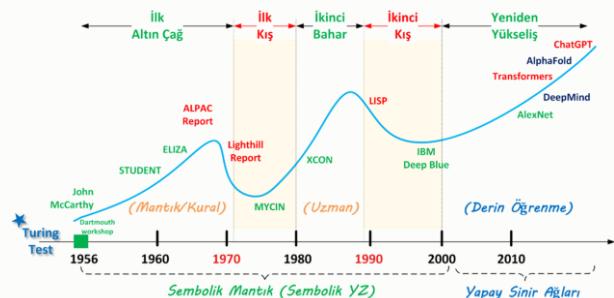
LLM'lerin Temel Sınırlamaları:

- Görünür başarısızlıklar —halüsinosyonlar,
- Gerçek nedensel muhakeme eksikliği,
- Eğitim verisi dışındaki temel görevlerde yetersizlik,
- **Istatistik model olması** (Non-deterministic System),
- Eğitim Maliyeti ve Hesaplama Limitleri: ““büyük daha iyidir””



32

“Winter is coming!”



Mevcut Göstergeler ("Çukur" Sinyalleri)

- Kurumsal Pilot Proje Başarısızlıkları: %95
- Yatırım Getirişi (ROI)
- Veri Kitliği ve Kirlenmesi (**model çöküşü** veya **veri zehirlenmesi**)
- Yanlış bilgilendirme, deepfake'ler, telif hakkı ihlali ...



33

“Winter is coming!”



Nobel Laureate Busts the AI Hype
MIT Sloan Management Review • 390K views • 5 months ago

Daron Acemoğlu

Önümüzdeki 10 yıl içinde bir işin alt görevlerinin ancak %5'inin dönüşmesi mümkün görünüyor, bu da ABD GSMH'sına en çok %1 katkı demek.



Yann LeCun: We Won't Reach AGI By Scaling Up LLMS
Alex Kantrowitz • 320K views • 5 months ago

Yann LeCun

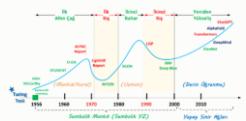
“büyük daha iyidir” iddiası yanlış. Daha çok veri ile AGI gelmeyecek.



“What's wrong with LLMs and what we should be building instead?”
valgrAI • 158K views • 2 years ago

Tom Dietterich

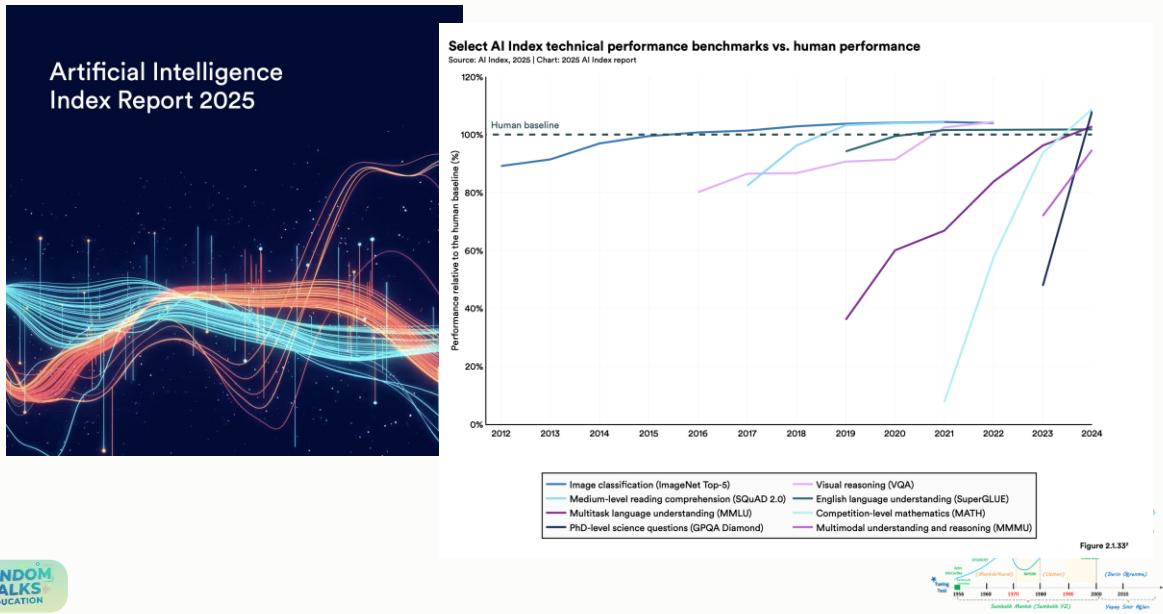
- LLM'lerle ilgili yanlış olan ne?
- AGI için, bunun yerine **ne yapmalyız?**



34

“AGI is coming!”

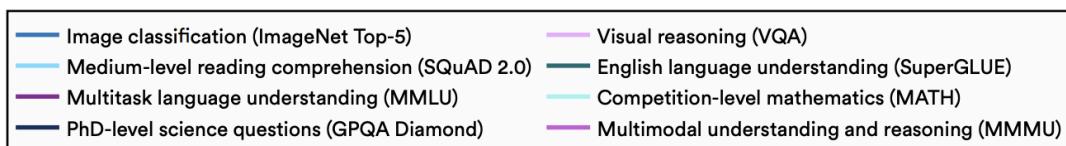
https://hai.stanford.edu/assets/files/hai_ai_index_report_2025.pdf



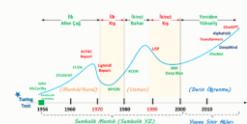
35

“AGI is coming!”

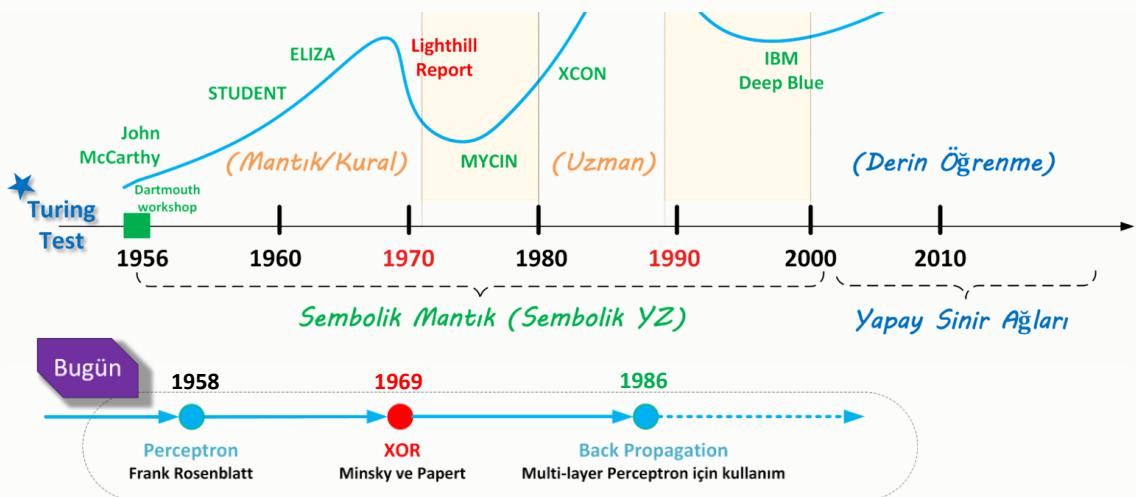
https://hai.stanford.edu/assets/files/hai_ai_index_report_2025.pdf



Syntaks düzeyinde desen/örgü tanımda insan düzeyinde başarılı!



36



37

Anlatım Akışı

1. Bölüm: Yapay Zeka alanında güncel durum

Geçmişten günümüze gelişmeler:

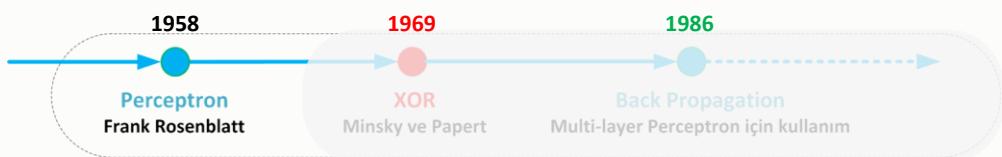
- Güncel durum
- Tarih tekerrür edecek mi?

2. Bölüm: Yapay Sinir Ağlarının Sıfırdan Keşfi

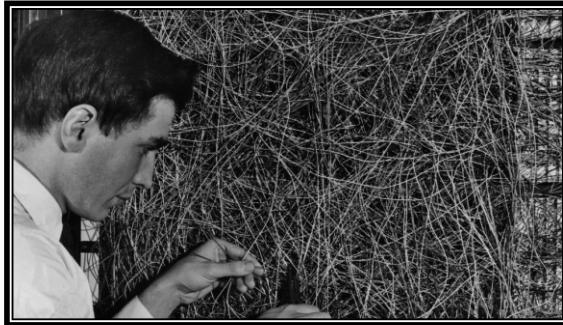
Çokça sorulan sorulara cevaplar:

- 1.Sınavda 12 saat çalışırsam kaç puan alırım?
- 2.Kaç saat çalışırsam dersten geçerim?
- 3.Hangi kaynaklardan sınava çalışayım?

38



"Sonunda orjinal fikirler üretmeye muktedir bir makine yapabildik"



1960 – Frank Rosenblatt
Mark 1 Perceptron Machine

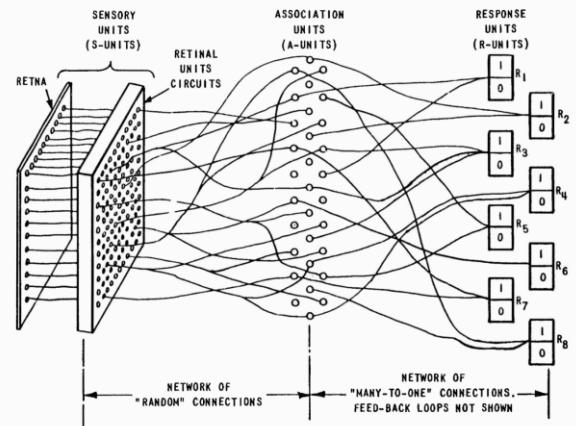
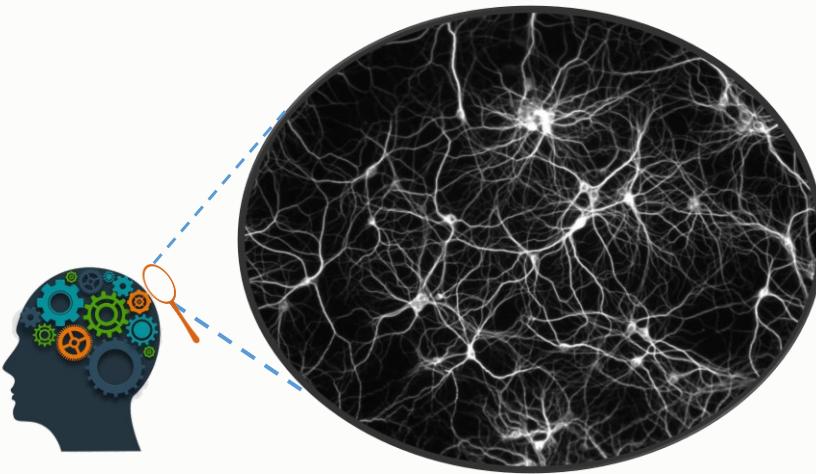


Figure I ORGANIZATION OF THE MARK I PERCEPTRON

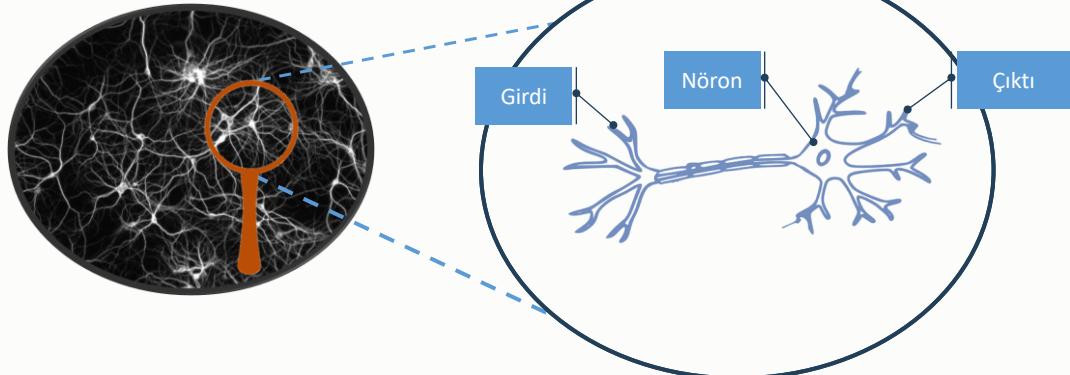
40

Aktivasyon halindeki nöronlar

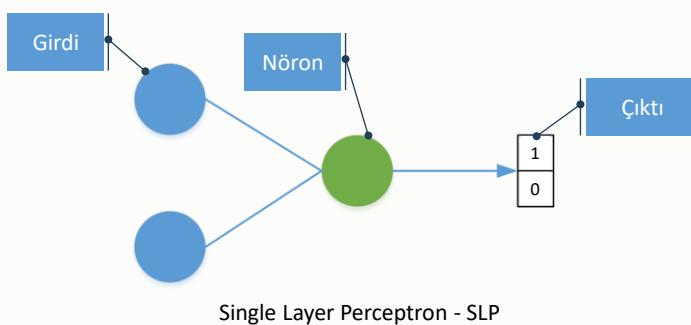
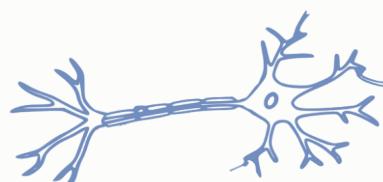


41

Aktivasyon halindeki nöronlar



42



Single Layer Perceptron - SLP



43

1960 – Frank Rosenblatt
Mark 1 Perceptron Machine

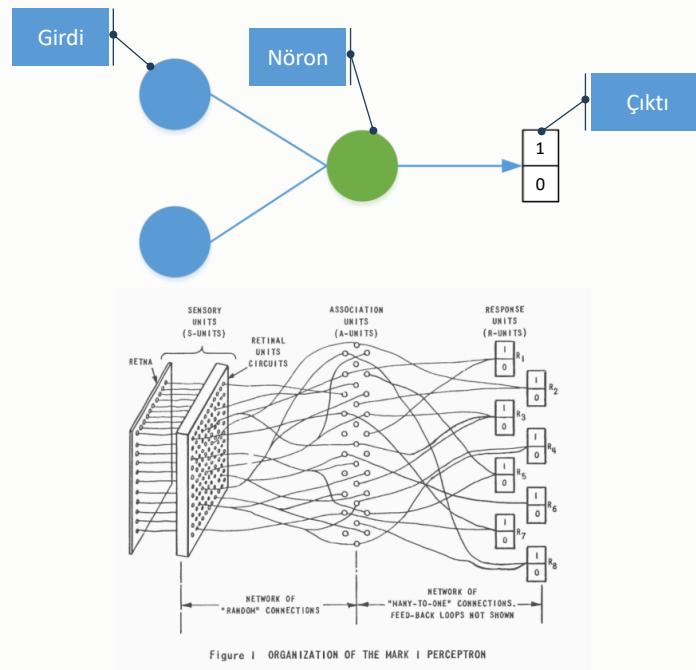


Figure I ORGANIZATION OF THE MARK I PERCEPTRON

44

1960 – Frank Rosenblatt
Mark 1 Perceptron Machine

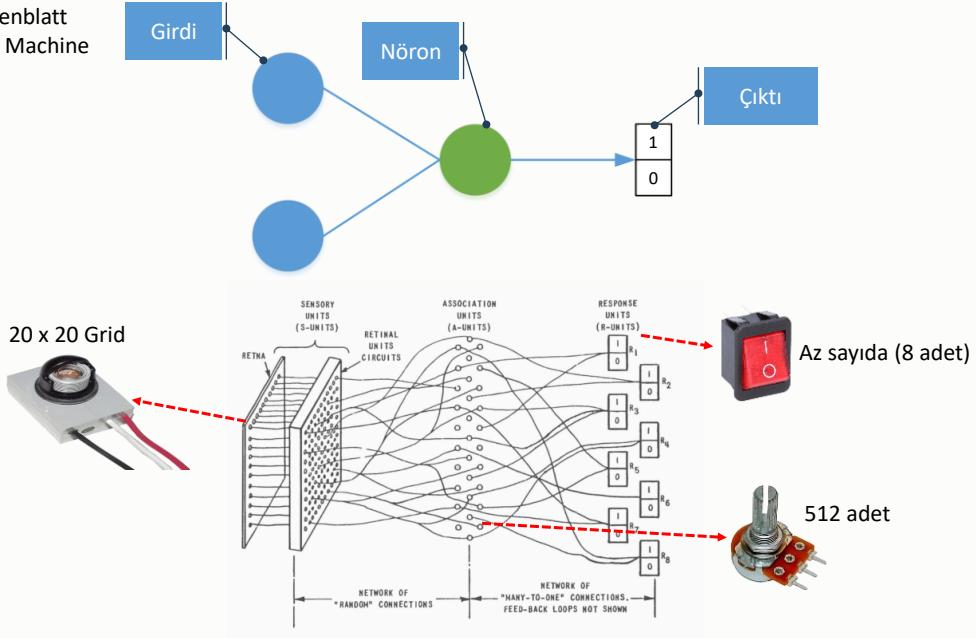
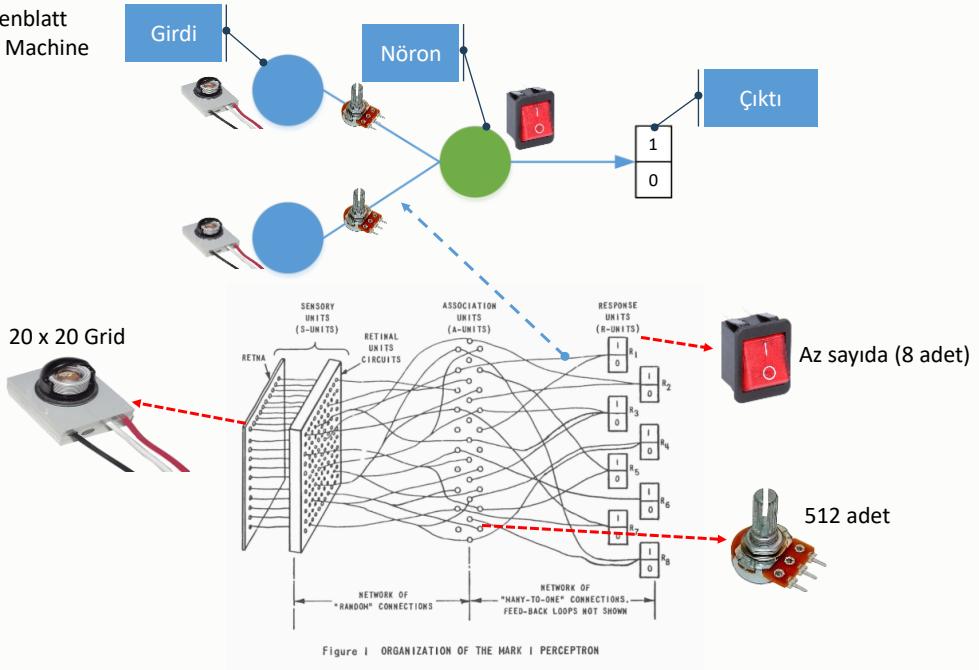


Figure I ORGANIZATION OF THE MARK I PERCEPTRON

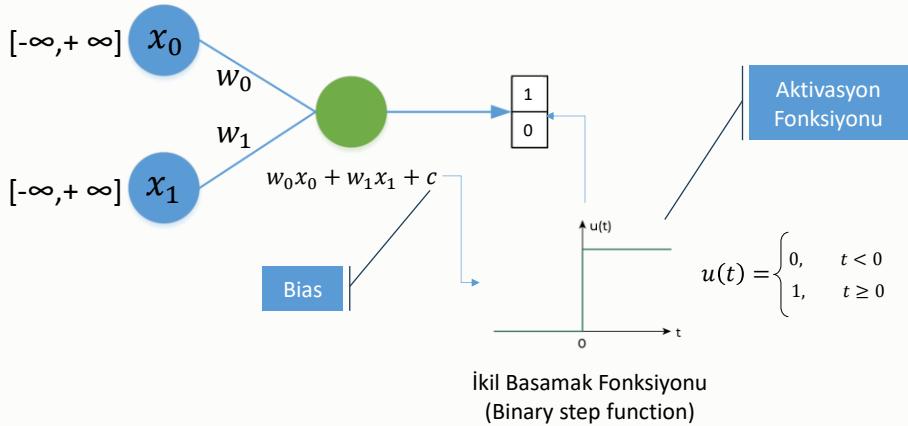
45

1960 – Frank Rosenblatt
Mark 1 Perceptron Machine



46

Rosenblatt's Single Layer Perceptron - SLP



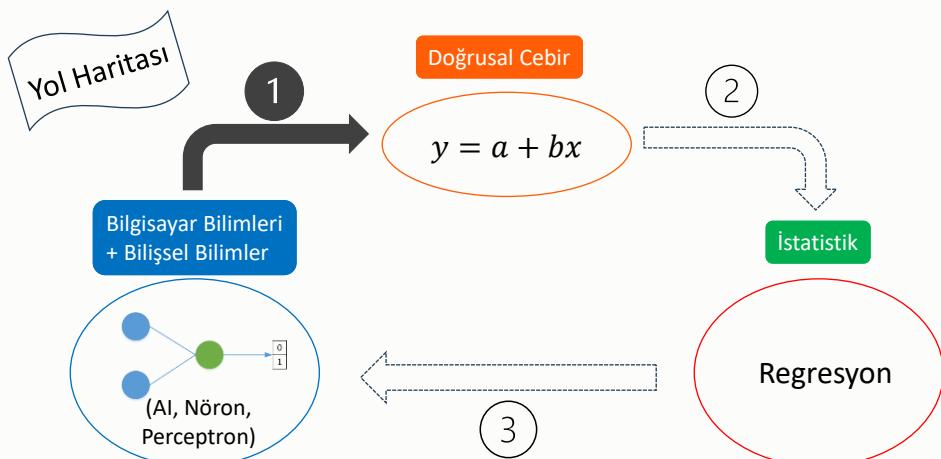
47

Rosenblatt's Single Layer Perceptron - SLP

$$\begin{aligned}
 & u(\underbrace{w_0x_0 + w_1x_1 + c}_{}) \\
 x_0 = 1 \rightarrow & w_0 + w_1x_1 + c \\
 \rightarrow & (w_0+c) + w_1x_1 \\
 \rightarrow & a + w_1x_1 \\
 & \boxed{y_1 = a + bx_1} \quad \text{Doğru denklemi}
 \end{aligned}$$



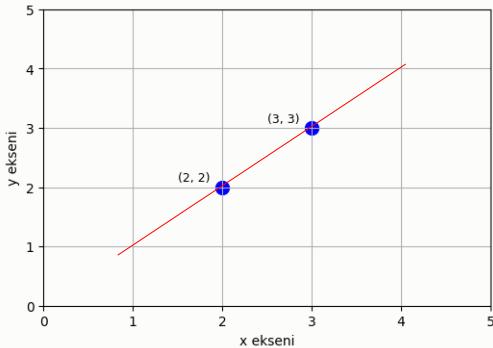
48



49

Doğrusal Cebir

$$y = a + bx$$



$$(2,2) \rightarrow 2 = a + 2b$$

$$(3,3) \rightarrow 3 = a + 3b$$

2 bilinmeyen (a ve b),
2 denklem

$$a = 0, b = 1$$

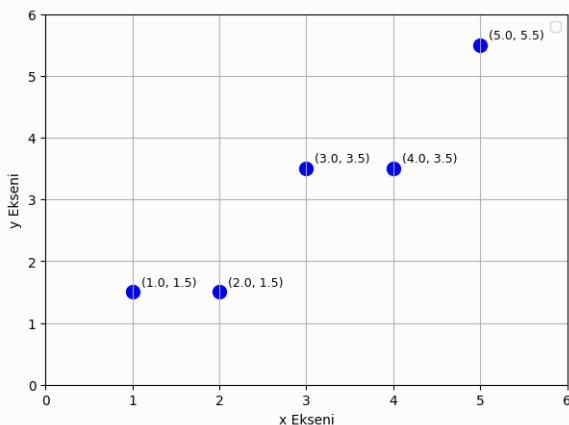
$$y = 0 + 1x$$



50

$$y = a + bx$$

İstatistik



$$1.5 = a + 1.0b$$

$$1.5 = a + 2.0b$$

$$3.5 = a + 3.0b$$

$$3.5 = a + 4.0b$$

$$5.5 = a + 5.0b$$

Elimizden hala 2 bilinmeyen,
ancak bu defa 5 denklem var.

$$a = ?, b = ?$$

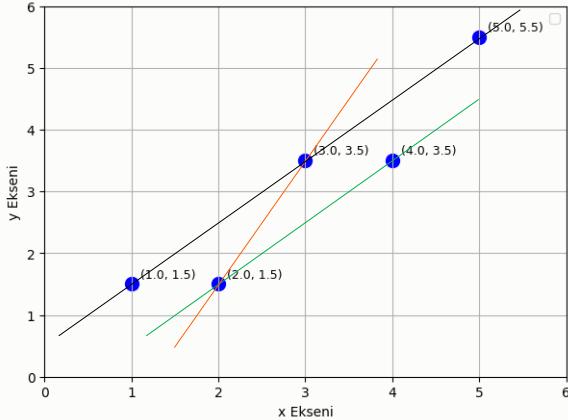


51

İstatistik

$$y = a + bx$$

Birden fazla doğru denklemi



$$1.5 = a + 1.0b$$

$$1.5 = a + 2.0b$$

$$3.5 = a + 3.0b$$

$$3.5 = a + 4.0b$$

$$5.5 = a + 5.0b$$

Elimizden halen 2 bilinmeyen,
ancak bu defa 5 denklem var.

$$a = ?, b = ?$$

Aşırın Tanımlı Sistem
(Over-determined Systems)

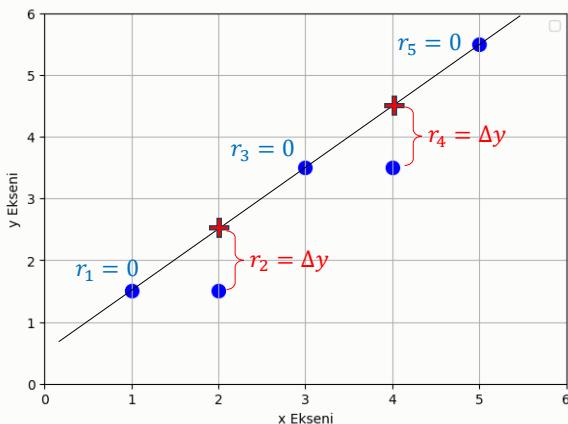


52

İstatistik

$$y = a + bx$$

$$\hat{y} = 0.5 + 1.0x$$



$$1.5 = a + 1.0b$$

$$1.5 = a + 2.0b$$

$$3.5 = a + 3.0b$$

$$3.5 = a + 4.0b$$

$$5.5 = a + 5.0b$$

Elimizden halen 2 bilinmeyen,
ancak bu defa 5 denklem var.

$$a = ?, b = ?$$

$$a = 0.5, b = 1.0$$

artık (residual)

$$r_i = (y_i - \hat{y}_i)$$

$$r_1 = (1.5 - 1.5) = 0$$

$$r_2 = (1.5 - 2.5) = -1$$

$$r_3 = (3.5 - 3.5) = 0$$

$$r_4 = (3.5 - 4.5) = -1$$

$$r_5 = (5.5 - 5.5) = 0$$



53

En Küçük Kareler Yöntemi (Least Squares)

İstatistik

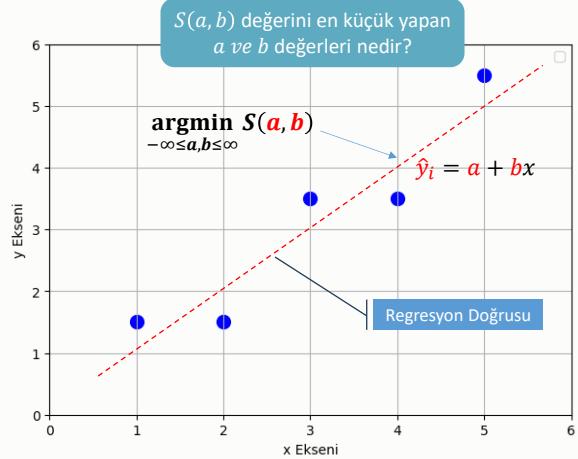
$$y = a + bx$$

Denklemler sistemi

$$\begin{aligned} y_1 &= a + bx_1 \\ y_2 &= a + bx_2 \\ y_3 &= a + bx_3 \\ \vdots & \\ y_n &= a + bx_n \end{aligned}$$

Hata Kareleri Toplamı (SSE)

$$\begin{aligned} S(a, b) &= \sum_{i=1}^n r_i^2 \\ &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 \end{aligned}$$



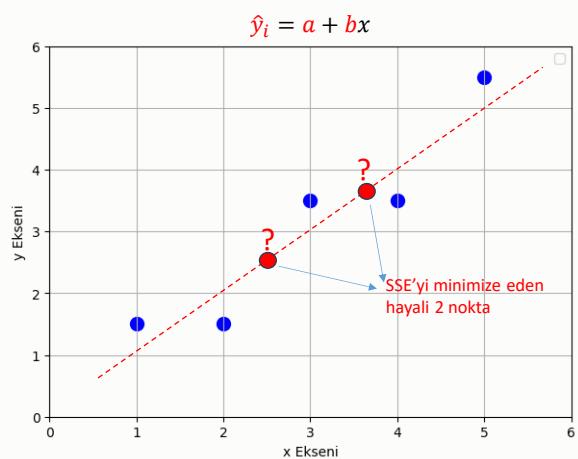
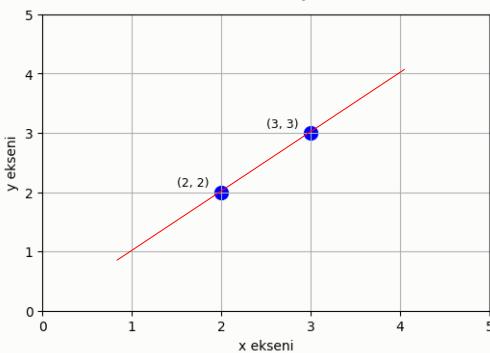
54

En Küçük Kareler Yöntemi (Least Squares)

İstatistik

$$y = a + bx$$

$$\begin{cases} (2, 2) \rightarrow 2 = a + 2b \\ (3, 3) \rightarrow 3 = a + 3b \end{cases} \quad \begin{cases} a = 0 \\ b = 1 \end{cases}$$



55

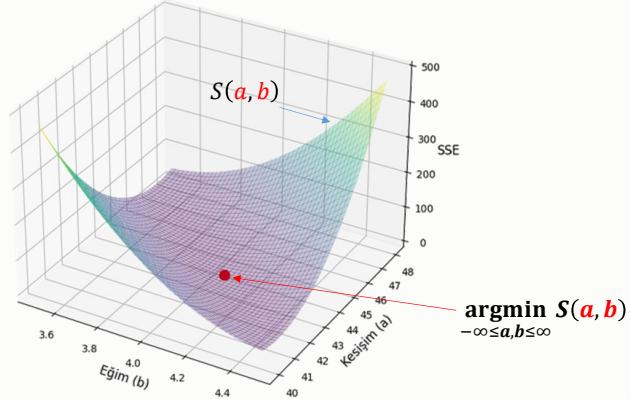
En Küçük Kareler Yöntemi (Least Squares)

İstatistik

$$y = a + bx$$

Hata Kareleri ToplAMI (SSE)

$$\begin{aligned} S(a, b) &= \sum_{i=1}^n r_i^2 \\ &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 \end{aligned}$$



56

En Küçük Kareler Yöntemi (Least Squares)

İstatistik

$$y = a + bx$$

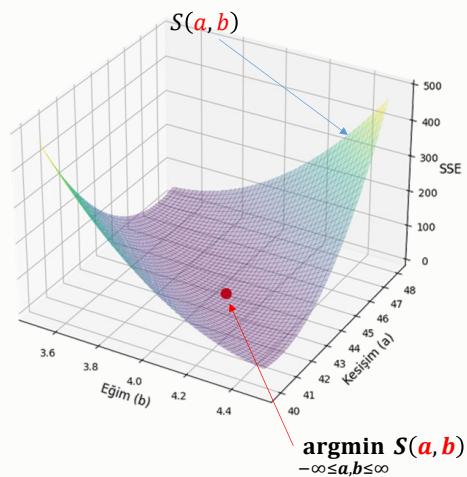
Hata Kareleri ToplAMI (SSE)

$$\begin{aligned} S(a, b) &= \sum_{i=1}^n r_i^2 \\ &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 \end{aligned}$$

Normal Denklemeleri

$$\left. \begin{aligned} \frac{\partial S(a, b)}{\partial b} &= 0 \\ \frac{\partial S(a, b)}{\partial a} &= 0 \end{aligned} \right\} \begin{matrix} 2 \text{ bilinmeyen} \\ 2 \text{ denklem} \end{matrix}$$

$$= \sum_{i=1}^n (y_i - (a + bx_i))^2$$



57

En Küçük Kareler Yöntemi (Least Squares)

İstatistik

$$y = a + bx$$

Hata Kareleri ToplAMI

$$S(a, b) = \sum_{i=1}^n r_i^2$$

$$= \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$= \sum_{i=1}^n (y_i - (a + bx_i))^2$$

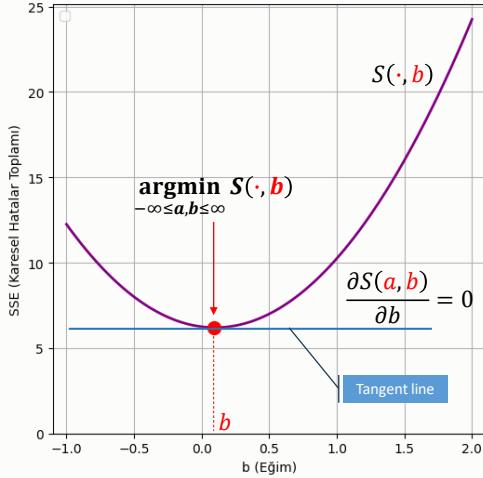
$$\frac{\partial S(a, b)}{\partial b} = 0$$



$$\sum_{i=1}^n 2(y_i - a - bx_i)(-x_i) = 0$$



$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$



58

En Küçük Kareler Yöntemi (Least Squares)

İstatistik

$$y = a + bx$$

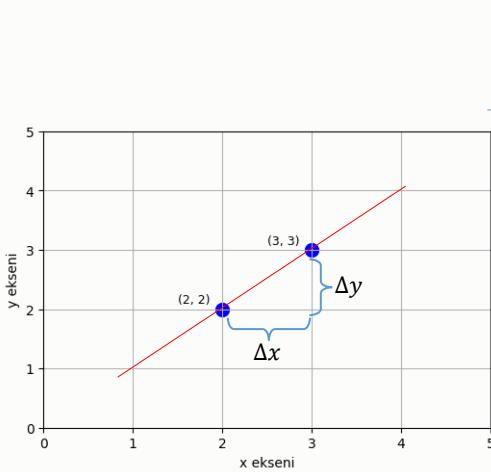
x değişimine karşılık gelen
 y değerlerindeki değişim

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{Cov(x, y)}{Var(x)}$$

x değerlerindeki değişim



59



$$y = a + bx$$

$$b = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_1 - x_2}$$

Doğrusal Cebir

Doğu denkleminde b parametresi (gradient):

x 'de 1 birim fark ($\Delta x = 1$) olduğunda,
 y 'de ne miktarda fark meydana geldiğini (Δy) ifade eder.



60

En Küçük Kareler Yöntemi (Least Squares)

$$y = a + bx$$

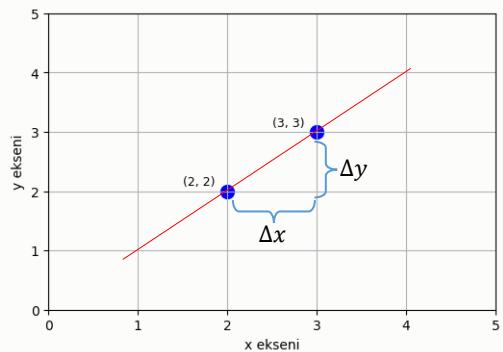
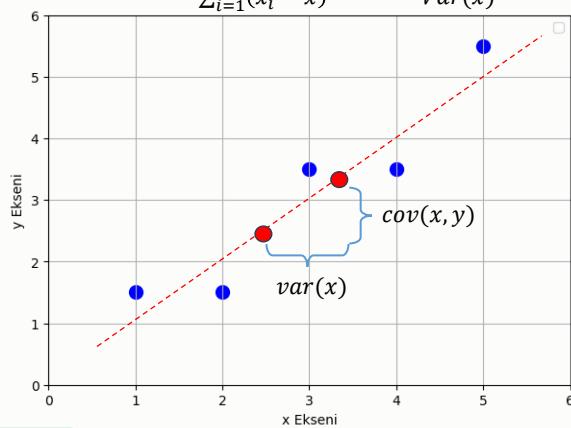
İstatistik

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{Cov(x, y)}{Var(x)}$$

≡

$$b = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_1 - x_2}$$

Doğrusal Cebir

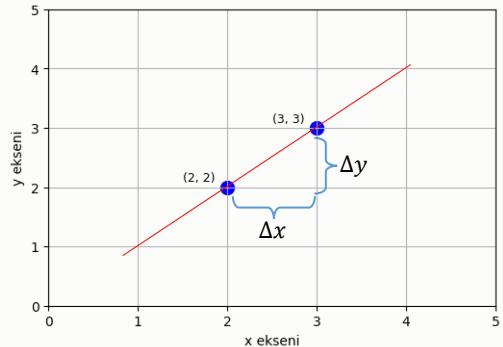
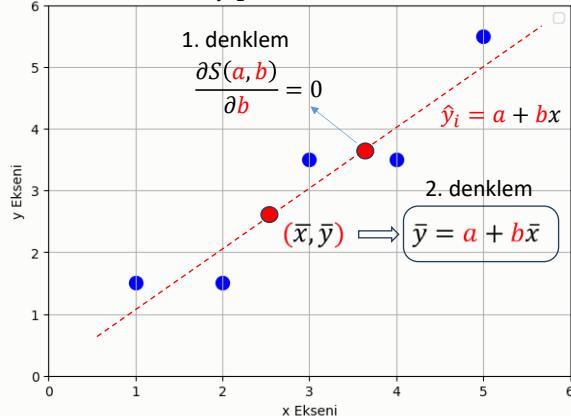


61

En Küçük Kareler Yöntemi (Least Squares)

$$y = a + bx$$

İstatistik $b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{Cov(x, y)}{Var(x)}$ \equiv $b = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_1 - x_2}$ Doğrusal Cebir



62

En Küçük Kareler Yöntemi (Least Squares)

İstatistik

$$y = a + bx$$

Hata Kareleri ToplAMI

$$S(a, b) = \sum_{i=1}^n r_i^2$$

$$= \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$= \sum_{i=1}^n (y_i - (a + bx_i))^2$$

Normal Denklemleri

$$\frac{\partial S(a, b)}{\partial b} = 0 \rightarrow b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad \text{1. denklem}$$

$$\frac{\partial S(a, b)}{\partial a} = 0 \rightarrow a = \bar{y} - b\bar{x} \Rightarrow \bar{y} = a + b\bar{x} \quad \text{2. denklem}$$

$\hat{y}_i = a + bx$ Regresyon Denklemi



63

“Hocam, sınavda 12 saat çalışırsam kaç puan alırım?”



64



<https://github.com/RandomTalksEducation/deep-learning-foundations/>



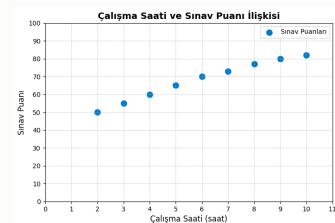
[Github repo](#)



65

“Hocam, sınava 12 saat çalışırsam kaç puan alırım?”

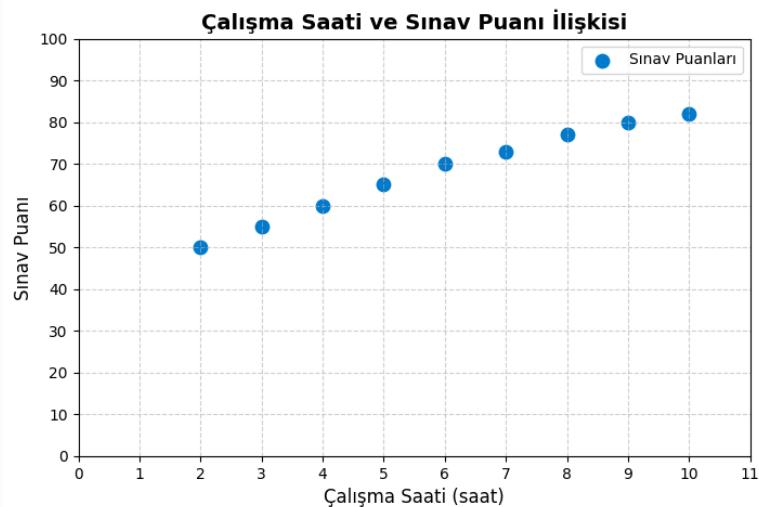
Ortalama Çalışma Saati (X)	Sınav Notu (Y)
2	50
3	55
4	60
5	65
6	73
7	77
8	80
10	82



66

“Hocam, sınava 12 saat çalışırsam kaç puan alırım?”

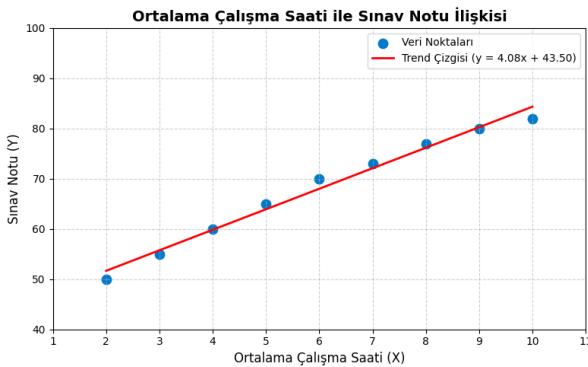
Ortalama Çalışma Saati (X)	Sınav Notu (Y)
2	50
3	55
4	60
5	65
6	73
7	77
8	80
10	82



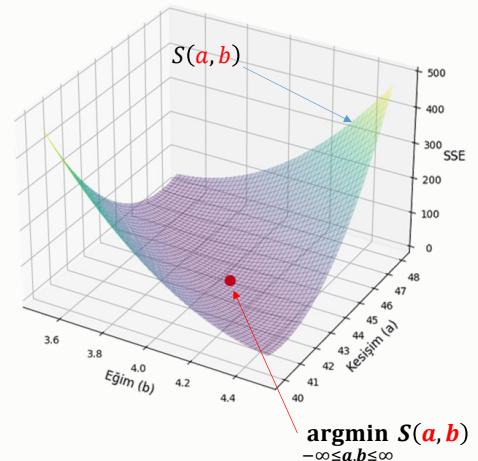
67



Doğrusal Regresyon ile eğilim (trend) çizgisi bulma

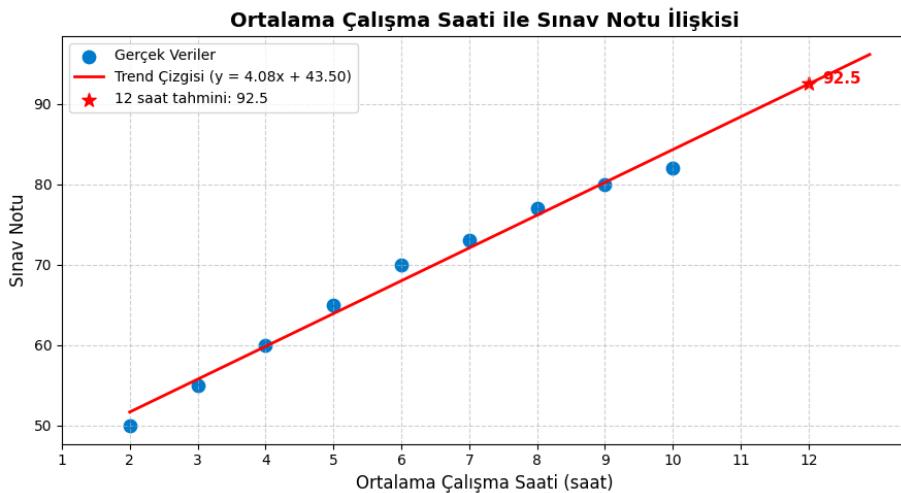


Hata kareleri toplamı grafiği (SSE)



68

12 Saat Çalışma için Not Tahimi



69

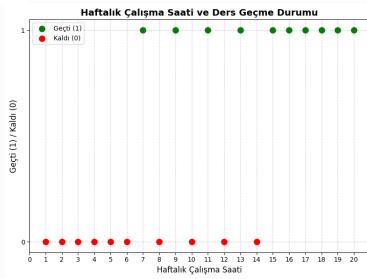
“Hocam, haftada kaç saat çalışırsam dersten geçerim?”



70

Geçmiş veriler

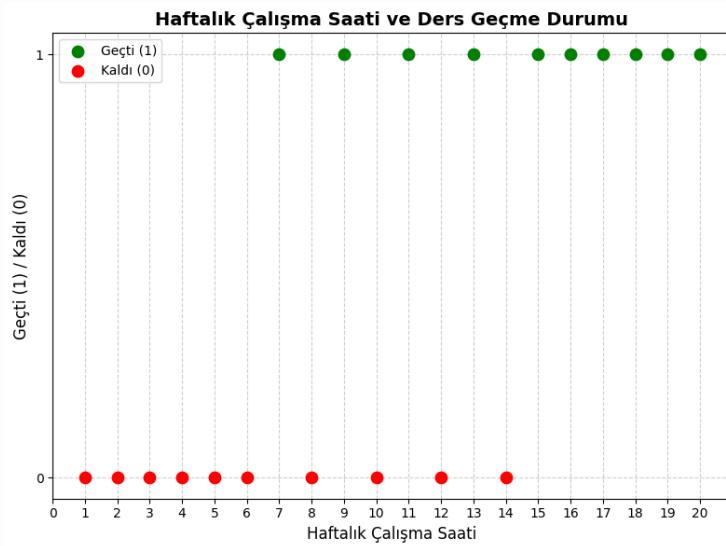
<input type="radio"/> Saat (x)	1	2	3	4	5	6	7	8	8	10
<input checked="" type="checkbox"/> Geçti (y)	0	0	0	0	0	0	1	0	1	0
<input type="radio"/> Saat (x)	11	12	13	14	15	16	17	18	19	20
<input checked="" type="checkbox"/> Geçti (y)	1	0	1	0	1	1	1	1	1	1



71

Geçmiş veriler

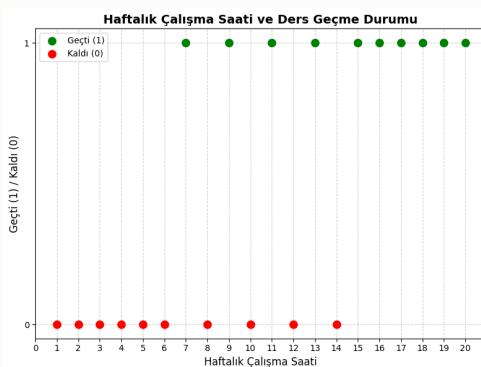
<input type="radio"/> Saat (x)	1	2	3	4	5	6	7	8	9	10
<input checked="" type="checkbox"/> Geçti (y)	0	0	0	0	0	0	1	0	1	0
<input type="radio"/> Saat (x)	11	12	13	14	15	16	17	18	19	20
<input checked="" type="checkbox"/> Geçti (y)	1	0	1	0	1	1	1	1	1	1



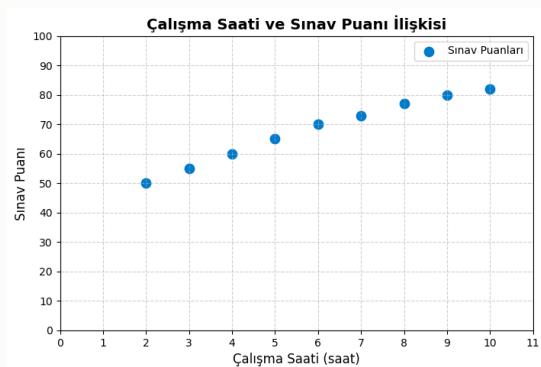
72

Çalışma saati vs Geçme (1)/Kalma (0) durumu için Yayılmış Çizgesi

Kategorik Değerler



Sürekli Değerler



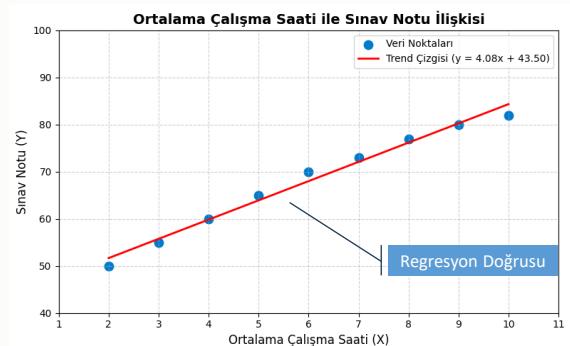
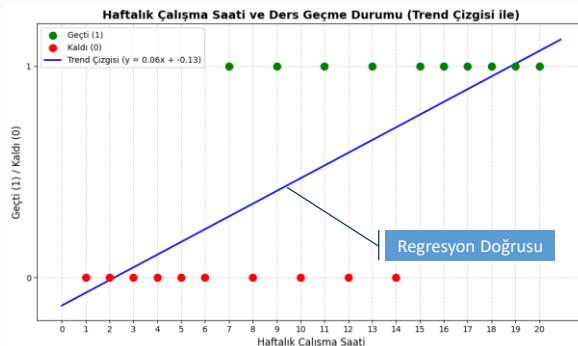
74

Çalışma saatı vs Geçme (1)/Kalma (0) durumu için Yayılmış Çizgesi

Kategorik Değerler

Regressyon Doğruları

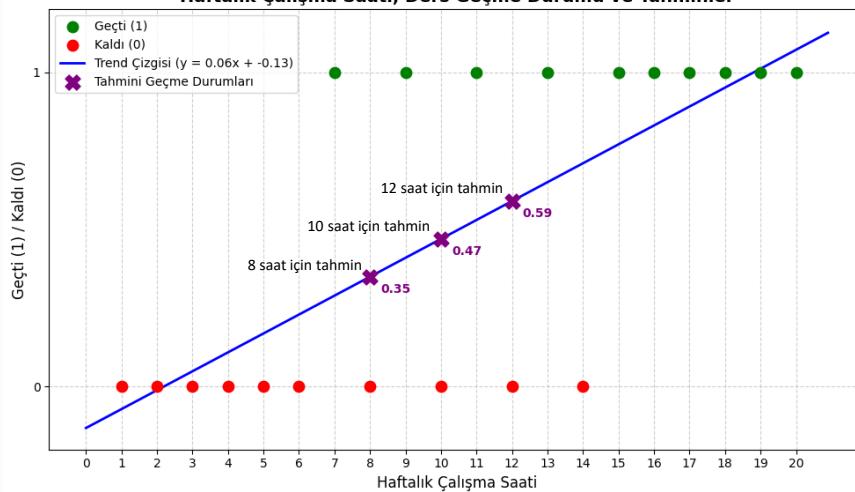
Sürekli Değerler



75

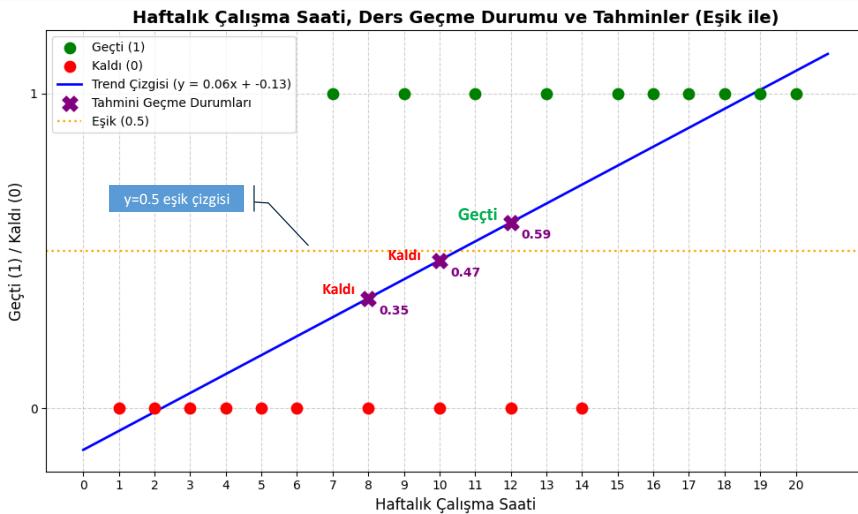
Çalışma saatı vs Geçme (1)/Kalma (0) durumu için Yayılmış Çizgesi

Haftalık Çalışma Saati, Ders Geçme Durumu ve Tahminler



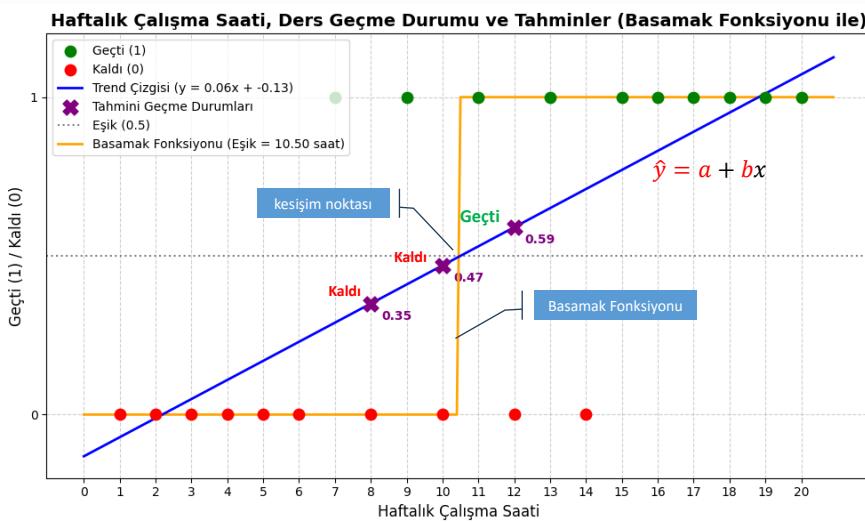
76

Çalışma saatı vs Geçme (1)/Kalma (0) durumu için Yayılmış Çizgesi



77

Çalışma saatı vs Geçme (1)/Kalma (0) durumu için Yayılmış Çizgesi

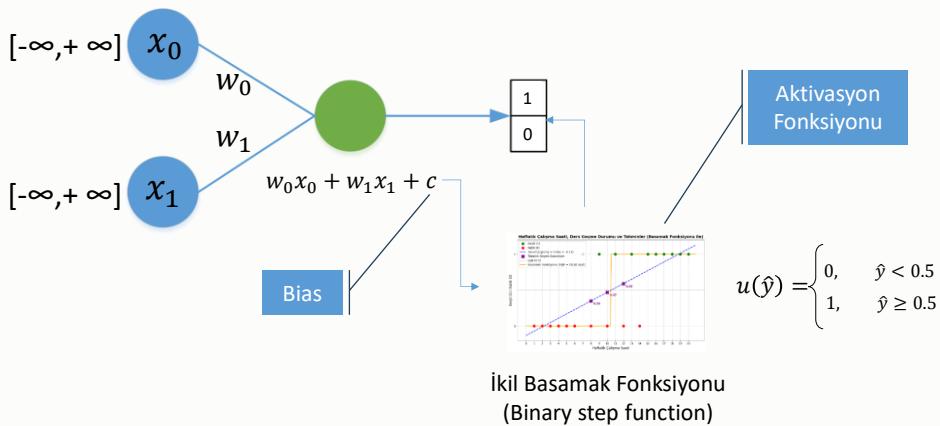


İkil Basamak Fonksiyonu
(Binary step function)

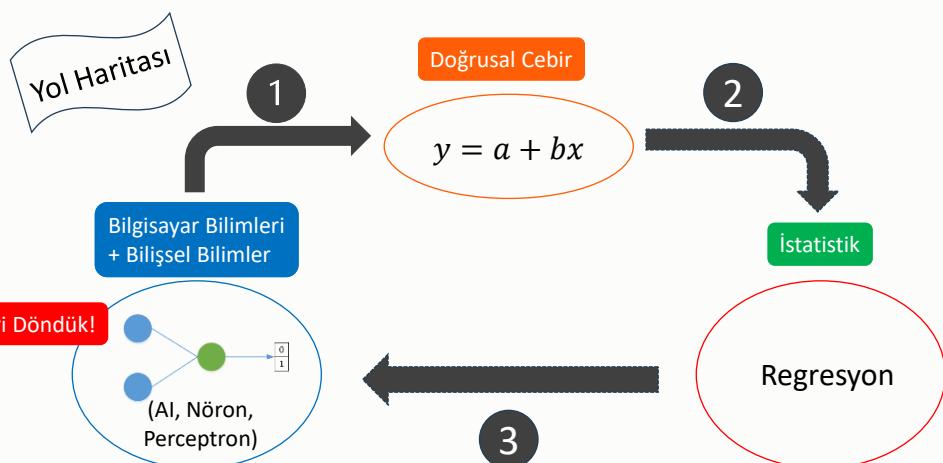
$$u(\hat{y}) = \begin{cases} 0, & \hat{y} < 0.5 \\ 1, & \hat{y} \geq 0.5 \end{cases}$$

78

Rosenblat'ın Single Layer Perceptron'u



79



80

Mark 1 Ayarları: "Makine Öğrenmesi!"

Mark 1'a
öğretilecek şekil

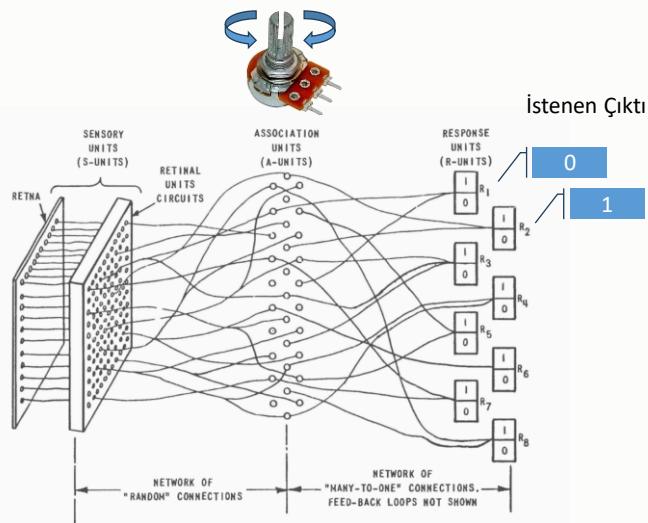


Figure I ORGANIZATION OF THE MARK I PERCEPTRON



81

Mark 1 Ayarları: "Makine Öğrenmesi!"

Mark 1'a
öğretilecek şekil

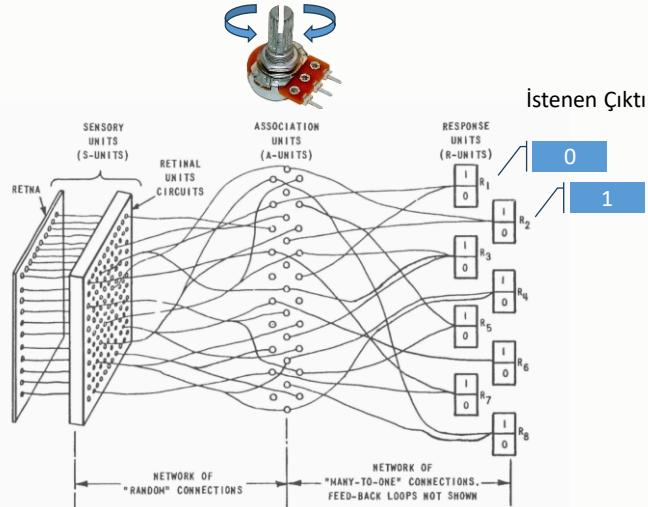


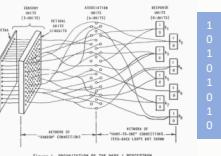
Figure I ORGANIZATION OF THE MARK I PERCEPTRON



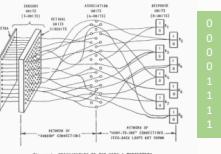
82

Mark 1 Ayarları: "Makine Öğrenmesi!"

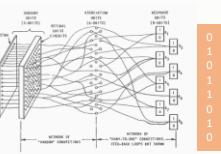
1. Örnek



2. Örnek



3. Örnek

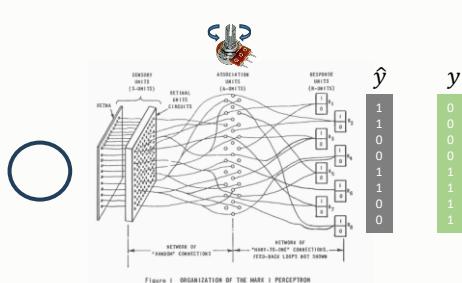


Elle adım adım yürütülen bir nevi
"Denetimli Öğrenim"
 (Supervise Learning)



83

Rosenblatt **neden** En Küçük Kareler Yöntemini Kullanmadı?



$$\text{Forward} \rightarrow v = a + bx \rightarrow \hat{y} = u(v)$$

$$a = a - \eta \nabla J(a) \quad (\hat{y} - y) \quad \text{Backward} \quad b = b - \eta \nabla J(b) \quad \text{Basamak fonksiyonu } u(v) \text{ nin türevi yok/tanımsız}$$

Zincir Kuralı (1676)

$$y = u(v), v = v(x)$$

$$\frac{dy}{dx} = \frac{dy}{dv} \cdot \frac{dv}{dx}$$



Gottfried Wilhelm Leibniz

Gradient Descent (1847)

$$\theta = \theta - \eta \nabla J(\theta)$$

$\nabla J(\theta)$: Tüm veri üzerinden grandyan

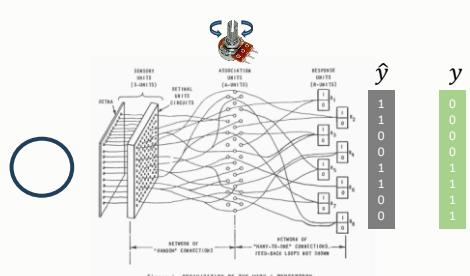
η : İniş oranı (learning rate)



Augustin-Louis Cauchy



Rosenblatt neden En Küçük Kareler Yöntemini Kullanmadı?



Forward $\rightarrow v = a + bx \rightarrow \hat{y} = u(v)$ 1960

$$\begin{aligned} a &= a - \eta \nabla J(a) \\ b &= b - \eta \nabla J(b) \end{aligned}$$

1974

$(\hat{y} - y)$

Backwards



Zincir Kuralı (1676)

$$y = u(v), v = v(x)$$



$$\frac{dy}{dx} = \frac{dy}{dv} \cdot \frac{dv}{dx}$$

Gottfried Wilhelm Leibniz

Gradient Descent (1847)

$$\theta = \theta - \eta \nabla J(\theta)$$

$\nabla J(\theta)$: Tüm veri üzerinden grandyan
 η : İniş oranı (learning rate)



Augustin-Louis Cauchy

backwards propagation of errors (1974)



PhD, Harvard



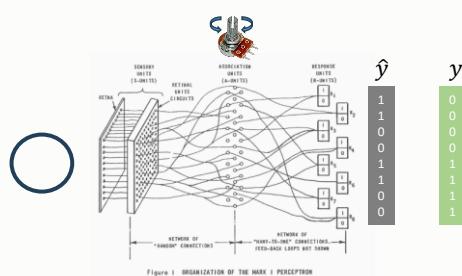
Reverse mode automatic differentiation (1970)



MSc, Helsinki

85

Rosenblatt neden En Küçük Kareler Yöntemini Kullanmadı?



Forward $\rightarrow v = a + bx \rightarrow \hat{y} = u(v)$ 1960

$$\begin{aligned} a &= a - \eta \nabla J(a) \\ b &= b - \eta \nabla J(b) \end{aligned}$$

?

$(\hat{y} - y)$

Backwards

1958



Zincir Kuralı (1676)

$$y = u(v), v = v(x)$$



Gottfried Wilhelm Leibniz

Gradient Descent (1847)

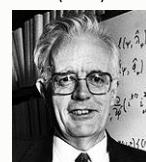
$$\theta = \theta - \eta \nabla J(\theta)$$

$\nabla J(\theta)$: Tüm veri üzerinden grandyan
 η : İniş oranı (learning rate)



Augustin-Louis Cauchy

Logistic Regression (1958)



Sir David Roxbee Cox

Logit

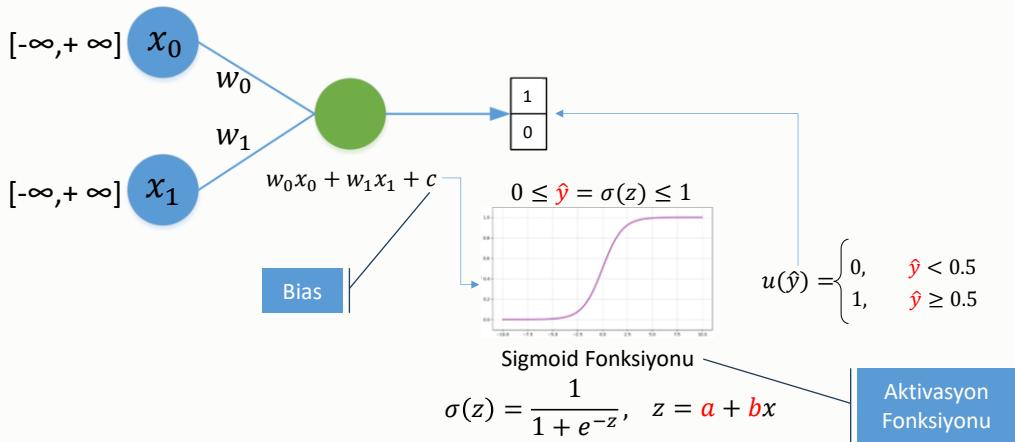
(1944)



Joseph Berkson

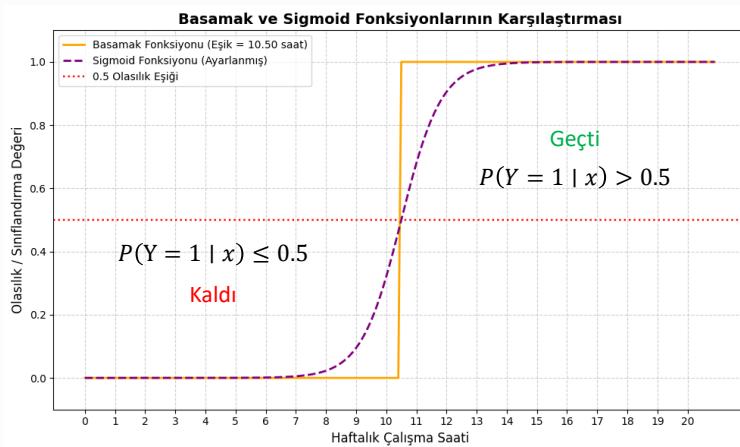
86

Logistic Regression ≡ Single Layer Perceptron



87

Logistic Regression



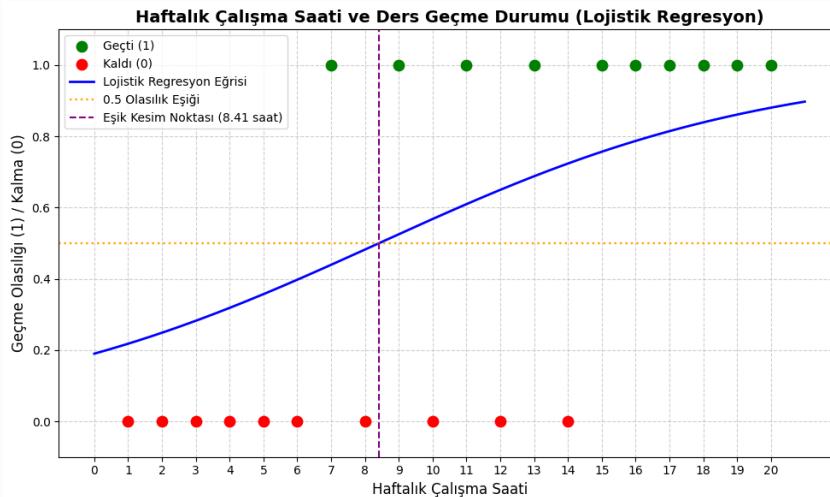
$$\hat{y} = \sigma(z) = P(Y = 1 | x) = \frac{1}{1 + e^{-(a+bx)}}$$

$$0 \leq \hat{y} \leq 1$$



88

Logistic Regression: Çalışma Saati Örneği



89

Gradient Descent

Doğrusal Regresyon

$$\hat{y} = a + bx$$

En Küçük Kareler (Least Squares)

$$S(a, b) = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Normal Denklemleri

$$\frac{\partial S(a, b)}{\partial b} = 0$$

$$\frac{\partial S(a, b)}{\partial a} = 0$$



90

Gradient Descent

Doğrusal Regresyon

$$\hat{y} = a + bx$$

En Küçük Kareler (Least Squares)

$$S(a, b) = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Normal Denklemleri

$$\frac{\partial S(a, b)}{\partial b} = 0$$

$$\frac{\partial S(a, b)}{\partial a} = 0$$

Notasyon Parametreyi Ekle/Cıkar

$$\frac{\partial S(a, b)}{\partial b} = 0 \rightarrow \nabla_b S(a, b) = 0 \rightarrow b = b - \nabla_b S(a, b)$$

$$\frac{\partial S(a, b)}{\partial a} = 0 \rightarrow \nabla_a S(a, b) = 0 \rightarrow a = a - \nabla_a S(a, b)$$



91

Gradient Descent: Doğrusal Regresyon

Doğrusal Regresyon

$$\hat{y} = a + bx$$

En Küçük Kareler (Least Squares)

$$S(a, b) = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Normal Denklemleri

$$\frac{\partial S(a, b)}{\partial b} = 0$$

$$\frac{\partial S(a, b)}{\partial a} = 0$$

Notasyon Parametreyi Ekle/Cıkar

$$\frac{\partial S(a, b)}{\partial b} = 0 \rightarrow \nabla_b S(a, b) = 0 \rightarrow b = b - \nabla_b S(a, b)$$

$$\frac{\partial S(a, b)}{\partial a} = 0 \rightarrow \nabla_a S(a, b) = 0 \rightarrow a = a - \nabla_a S(a, b)$$

Gradient Descent (1847)

$$\theta = \theta - \eta \nabla J(\theta)$$



Augustin-Louis Cauchy



$$b = b - \eta \nabla_b S(a, b)$$

$$a = a - \eta \nabla_a S(a, b)$$

$$0 < \eta < 1$$



92

Gradient Descent: Doğrusal Regresyon

Doğrusal Regresyon

$$\hat{y} = a + bx$$

En Küçük Kareler (Least Squares)

$$S(a, b) = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Normal Denklemleri

$$\frac{\partial S(a, b)}{\partial b} = 0$$

$$\frac{\partial S(a, b)}{\partial a} = 0$$

Notasyon

$$\frac{\partial S(a, b)}{\partial b} = 0 \rightarrow \nabla_b S(a, b) = 0 \rightarrow b = b - \nabla_b S(a, b)$$

Parametreyi Ekle/Cıkar

$$\frac{\partial S(a, b)}{\partial a} = 0 \rightarrow \nabla_a S(a, b) = 0 \rightarrow a = a - \nabla_a S(a, b)$$

Gradient Descent (1847)



Augustin-Louis Cauchy

$$\theta = \theta - \eta \nabla J(\theta)$$

(a,b) için ilk değerler ne olursa olsun $\eta = 1$ olduğunda, sonuç SSE'yi minimize eden (a,b) değerleri olur.

$$b = b - \eta \nabla_b S(a, b)$$

$$a = a - \eta \nabla_a S(a, b)$$

$$0 < \eta < 1$$

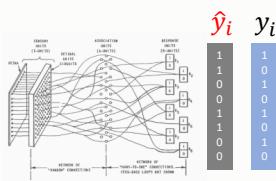


93

Mark 1 Ayarları: "Makine Öğrenmesi!"

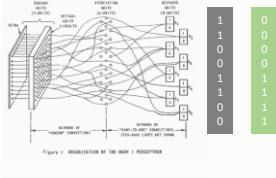
Her veri noktası için a ve b güncellemesi

1. Örnek



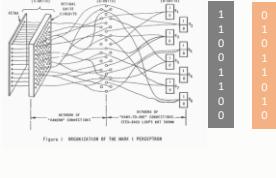
\hat{y}_i	y_i
1	1
1	0
0	1
0	1
1	0
1	1
0	0
0	1

2. Örnek



\hat{y}_i	y_i
0	0
0	0
0	0
1	1
1	1
1	1
0	0
0	1

3. Örnek



\hat{y}_i	y_i
1	0
1	1
0	1
0	1
1	0
1	1
0	0
0	1

Rasgele değerler: $a_0 = 21231$ $b_0 = 33321$

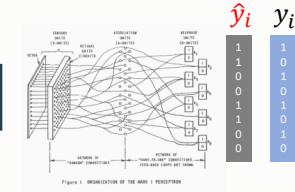
Forward $\rightarrow \hat{y}_1 = a_0 + b_0 x_1 \rightarrow (\hat{y}_1 - y_1)$
 $b_1 = b_0 - \eta \nabla_b S(a, b) \leftarrow \eta = 0.01 \leftarrow (\hat{y}_1 - y_1)$ Backward



94

Mark 1 Ayarları: "Makine Öğrenmesi!"

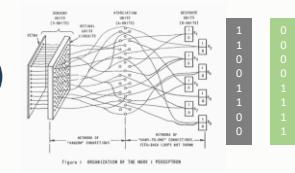
1. Örnek



$\hat{y}_i \quad y_i$

1	0
1	0
0	1
0	1
1	0
1	0
0	1
0	1
0	0

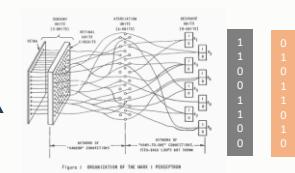
2. Örnek



$\hat{y}_i \quad y_i$

1	0
1	0
0	0
0	0
1	1
1	1
0	1
1	1
0	0

3. Örnek



$\hat{y}_i \quad y_i$

1	0
1	0
0	1
0	1
1	0
1	0
0	1
1	0
0	1

Her veri noktası için a ve b güncellemesi

Rasgele değerler: $a_0 = 21231 \quad b_0 = 33321$

$$\begin{array}{l} \text{Forward} \\ \boxed{\square} \end{array} \quad \hat{y}_1 = a_0 + b_0 x_1 \quad \eta = 0.01 \quad (\hat{y}_1 - y_1)$$

$$b_1 = b_0 - \eta \nabla_b S(a, b) \quad \text{Backward}$$

$$\begin{array}{l} \text{Forward} \\ \boxed{\circ} \end{array} \quad \hat{y}_2 = a_1 + b_1 x_1 \quad \eta = 0.01 \quad (\hat{y}_2 - y_2)$$

$$b_2 = b_1 - \eta \nabla_b S(a, b) \quad \text{Backward}$$

Her veri noktası için a ve b güncellemesi

Rasgele değerler: $a_0 = 21231 \quad b_0 = 33321$

$$\begin{array}{l} \text{Forward} \\ \boxed{\square} \end{array} \quad \hat{y}_1 = a_0 + b_0 x_1 \quad \eta = 0.01 \quad (\hat{y}_1 - y_1)$$

$$b_1 = b_0 - \eta \nabla_b S(a, b) \quad \text{Backward}$$

$$\begin{array}{l} \text{Forward} \\ \boxed{\circ} \end{array} \quad \hat{y}_2 = a_1 + b_1 x_1 \quad \eta = 0.01 \quad (\hat{y}_2 - y_2)$$

$$b_2 = b_1 - \eta \nabla_b S(a, b) \quad \text{Backward}$$

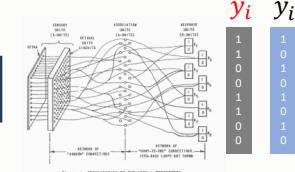
$$\begin{array}{l} \text{Forward} \\ \boxed{\triangle} \end{array} \quad \hat{y}_3 = a_2 + b_2 x_1 \quad \eta = 0.01 \quad (\hat{y}_3 - y_3)$$

$$b_3 = b_2 - \eta \nabla_b S(a, b) \quad \text{Backward}$$

95

Mark 1 Ayarları: "Makine Öğrenmesi!"

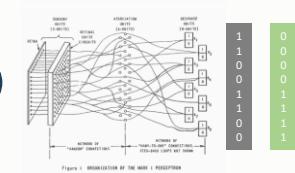
1. Örnek



$\hat{y}_i \quad y_i$

1	1
1	1
0	0
0	1
1	0
1	0
0	1
0	1
0	0

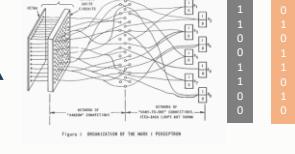
2. Örnek



$\hat{y}_i \quad y_i$

1	1
1	1
0	0
0	1
1	1
1	1
0	1
1	1
0	0

3. Örnek



$\hat{y}_i \quad y_i$

1	0
1	0
0	1
0	1
1	0
1	0
0	1
1	0
0	1



96

47

Gradient Descent: Jojistik Regresyon

Lojistik Regresyon

$$\hat{y} = \sigma(z) = P(Y = 1 | x) = \frac{1}{1 + e^{-(a+bx)}}$$

Hata/Maliyet Fonksiyonu:

İkil Çapraz Entropi (Binary Cross-Entropy)

$$L_{CE}(a, b) = -\frac{1}{n} \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

Tek bir veri noktası (x_i, y_i) için:

$$\begin{aligned} \text{gerçek değer } y_i = 1 &\rightarrow L_{CE}(y_i, \hat{y}_i) = -y_i \log(\hat{y}_i) \\ &= -\log(\hat{y}_i) \end{aligned}$$

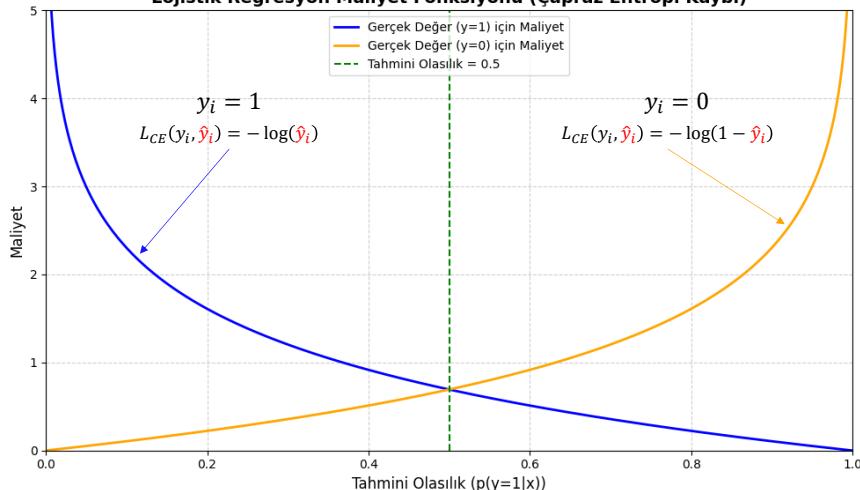
$$\begin{aligned} \text{gerçek değer } y_i = 0 &\rightarrow L_{CE}(y_i, \hat{y}_i) = -(1 - y_i) \log(1 - \hat{y}_i) \\ &= -\log(1 - \hat{y}_i) \end{aligned}$$



97

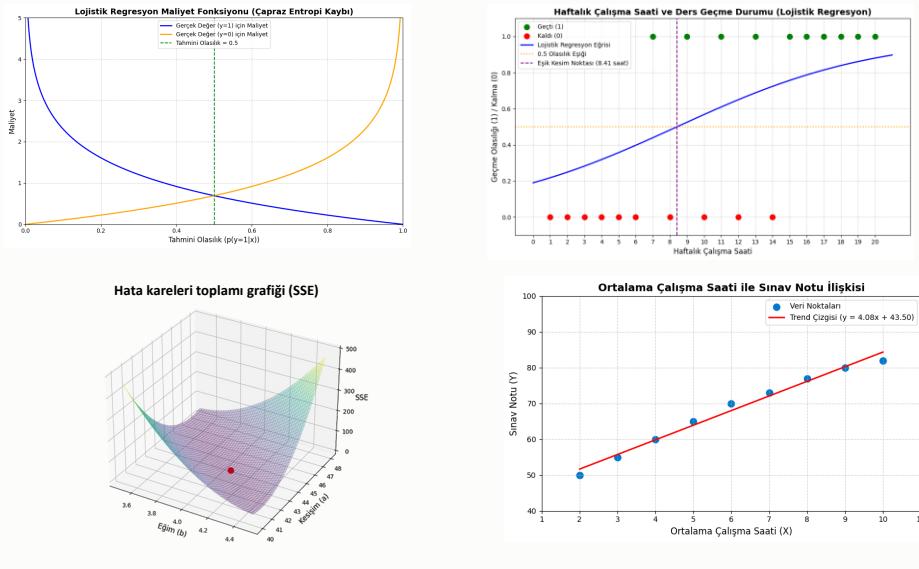
Lojistik Regresyon: Hata/Maliyet Fonksiyonu

Lojistik Regresyon Maliyet Fonksiyonu (Çapraz Entropi Kaybı)



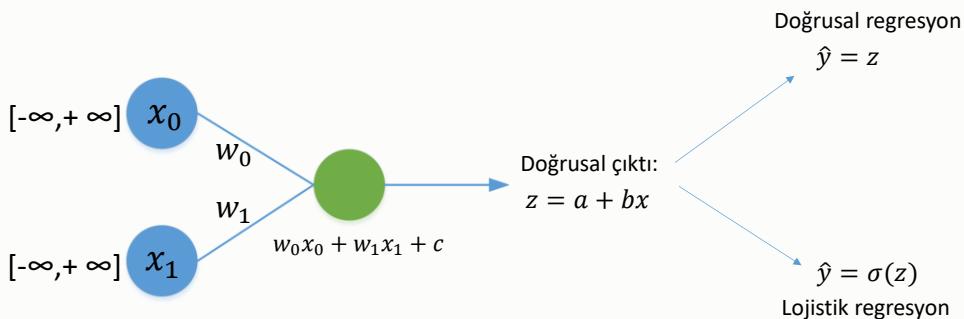
98

Lojistik Regresyon: Hata/Maliyet Fonksiyonu



99

Gradyanların Denkliği



101

Gradyanların Denkliği

Doğrusal çıktı: $z = a + bx$

Doğrusal regresyon

$$\hat{y} = z$$

Lojistik regresyon

$$\hat{y} = \sigma(z)$$

Tek bir veri noktası (x_i, y_i) için:

$$S(x_i, y_i) = \frac{1}{2}(y_i - \hat{y}_i)^2$$

$$L_{CE}(x_i, y_i) = -[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

Doğrusal Çıktıyla göre Gradyan:

$$\frac{\partial S(x_i, y_i)}{\partial z_i} = (\hat{y}_i - y_i)$$

$$\frac{\partial L_{CE}(x_i, y_i)}{\partial z_i} = (\hat{y}_i - y_i)$$

Doğrusal Model



Zincir Kuralı (1676)



$y = u(v), v = v(x)$

$$\frac{dy}{dx} = \frac{dy}{dv} \cdot \frac{dv}{dx}$$

Gottfried Wilhelm Leibniz

102

Sonuç

Single Layer Perceptron
Doğrusal Bir Model

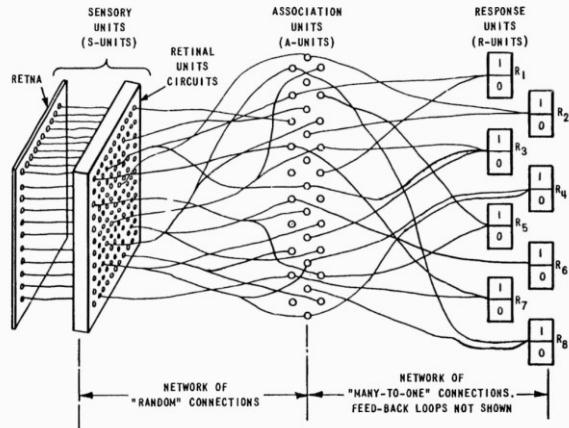


Figure 1 ORGANIZATION OF THE MARK I PERCEPTRON



103

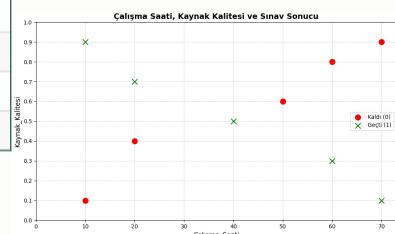
“Hocam, hangi kaynaklardan sınava çalışayım?”



104

Geçmiş Sınavlara ait Eldeki Veriler

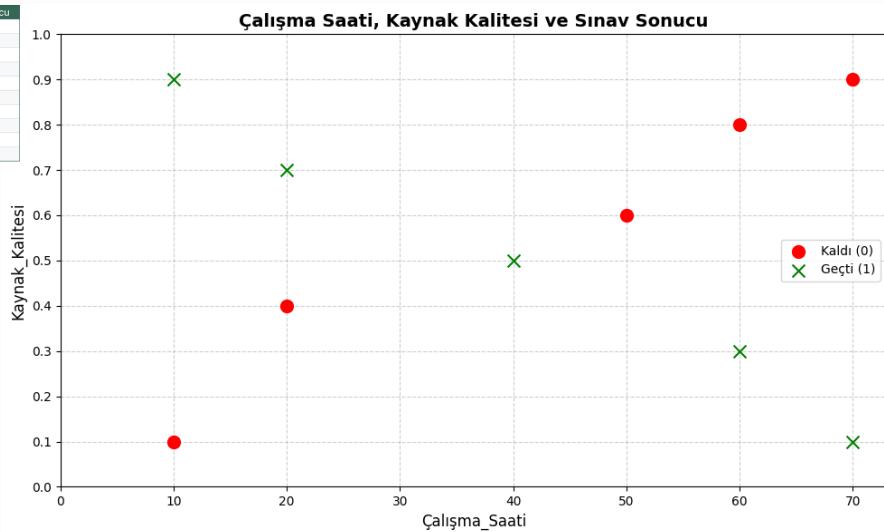
Çalışma Saati	Kaynak Kalitesi	Sınav Sonucu
10	0.1	0
10	0.9	1
20	0.4	0
20	0.7	1
40	0.5	1
50	0.6	0
60	0.3	1
60	0.8	0
70	0.1	1
70	0.9	0



106

Geçmiş Sınavlara ait Eldeki Veriler

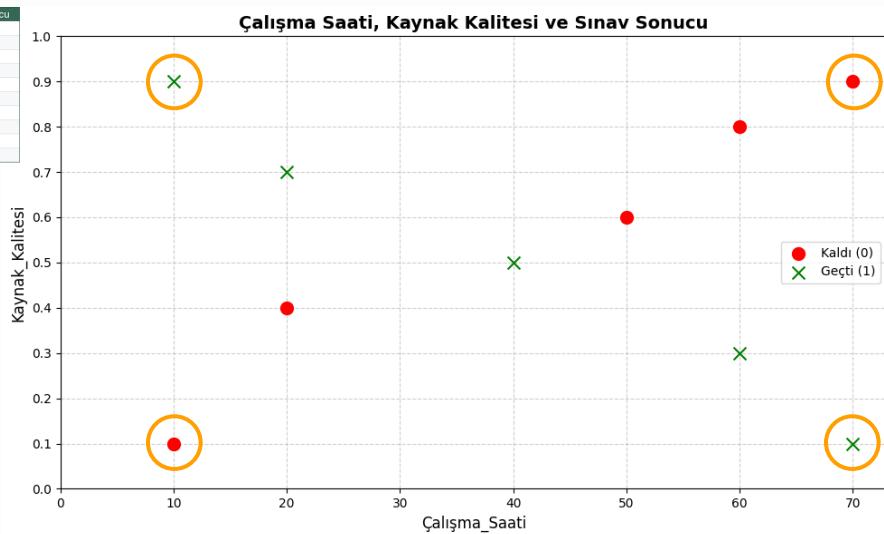
Çalışma Saati	Kaynak Kalitesi	Sınav Sonucu
10	0.1	0
10	0.9	1
20	0.4	0
20	0.7	1
40	0.5	1
50	0.6	0
60	0.3	1
60	0.8	0
70	0.1	1
70	0.9	0



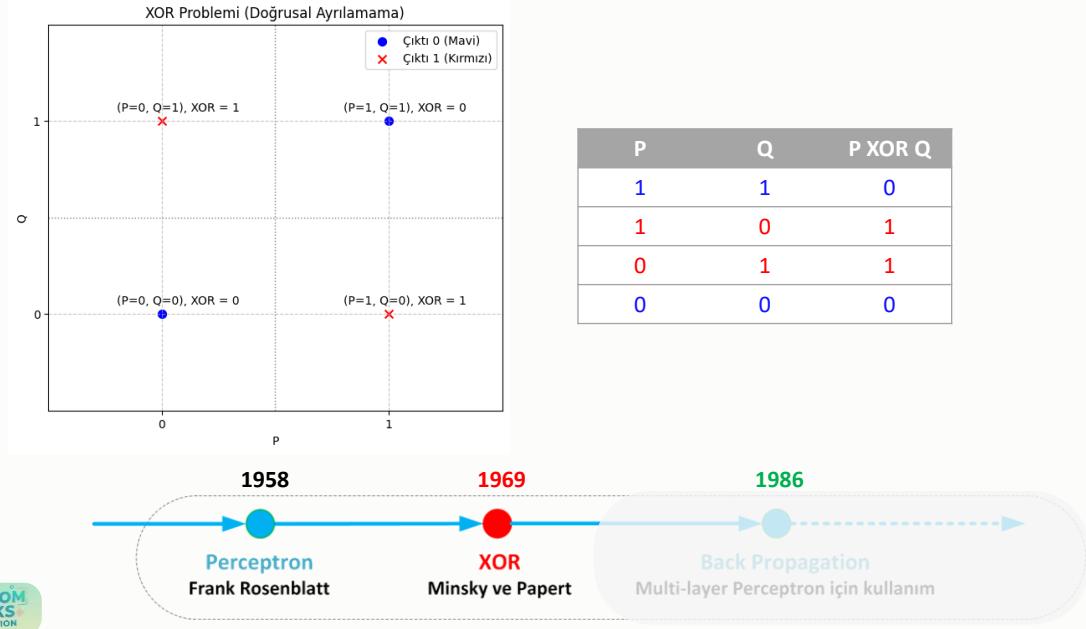
107

Geçmiş Sınavlara ait Eldeki Veriler

Çalışma Saati	Kaynak Kalitesi	Sınav Sonucu
10	0.1	0
10	0.9	1
20	0.4	0
20	0.7	1
40	0.5	1
50	0.6	0
60	0.3	1
60	0.8	0
70	0.1	1
70	0.9	0

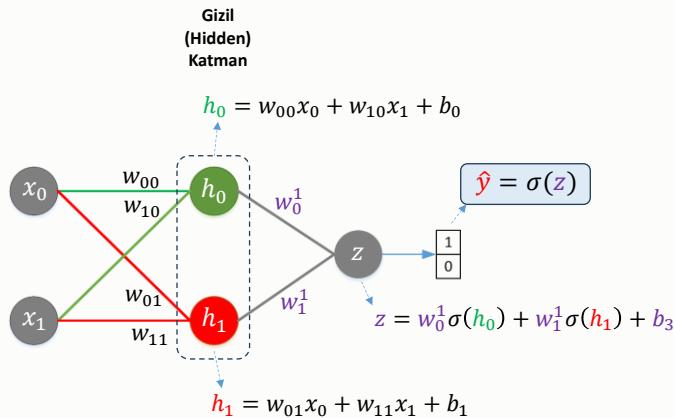


108



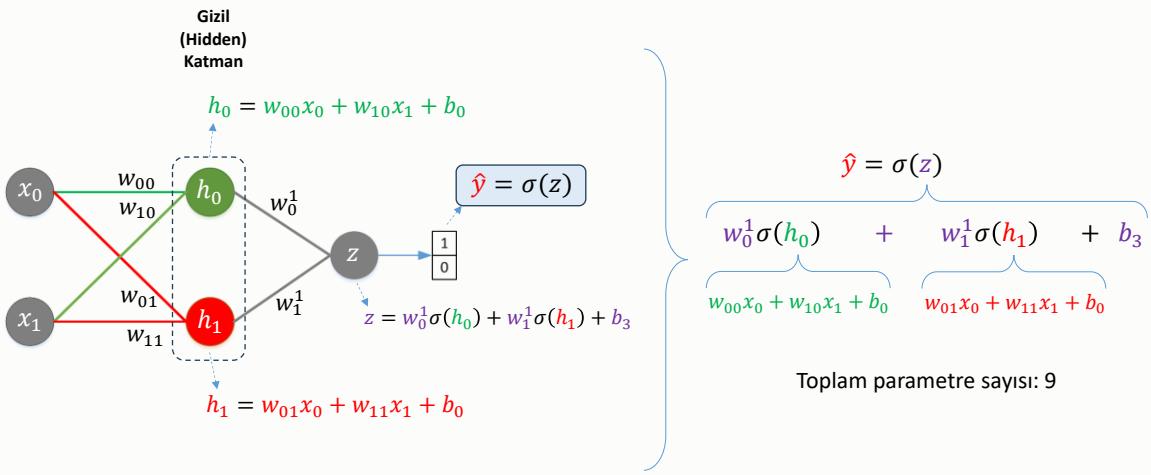
109

Çok Katmanlı Algılayıcı (Multilayer Perceptron - MLP)



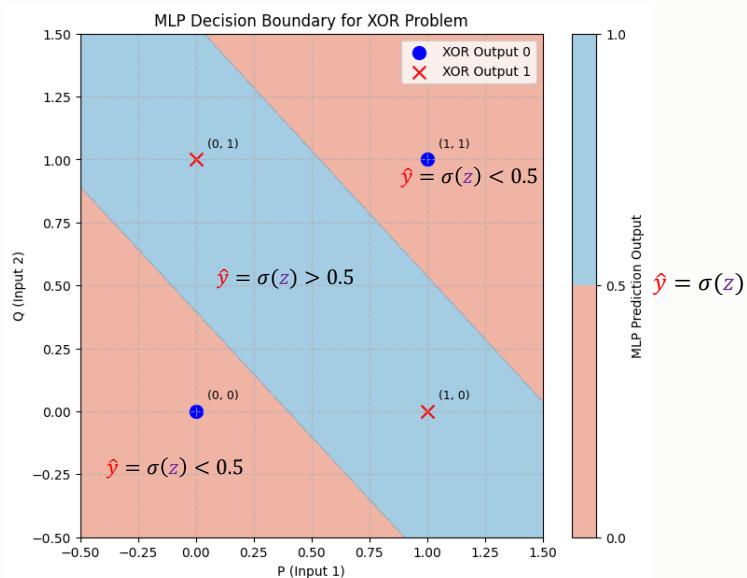
110

Çok Katmanlı Algılayıcı (Multilayer Perceptron - MLP)



111

MLP ile
XOR Problemi
çözümü

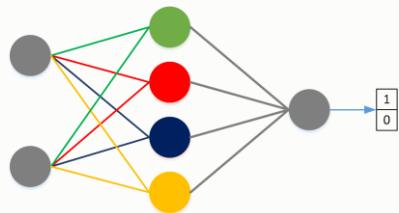


112

MLPs

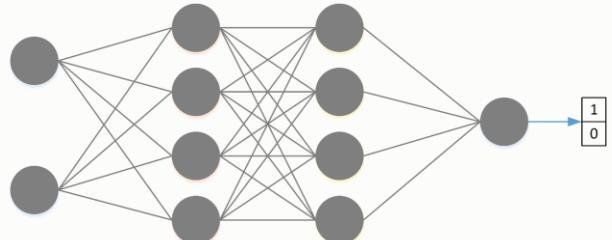
Derin Öğrenme
(Deep Learning)

Gizil Katmanda Node Sayısı Artırılabilir



Toplam parametre sayısı: 15

Gizil Katman Sayısı Artırılabilir

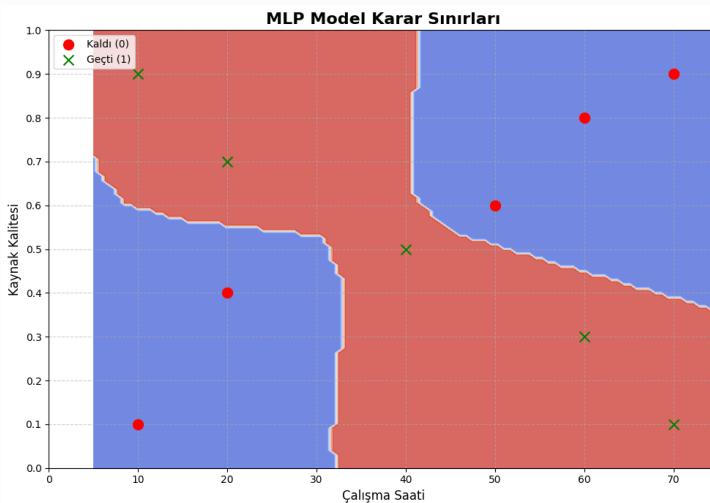


Toplam parametre sayısı: 25



113

Çalışma Saati – Kaynak Kalitesi için MLP sonuçları



MLP Özellikleri

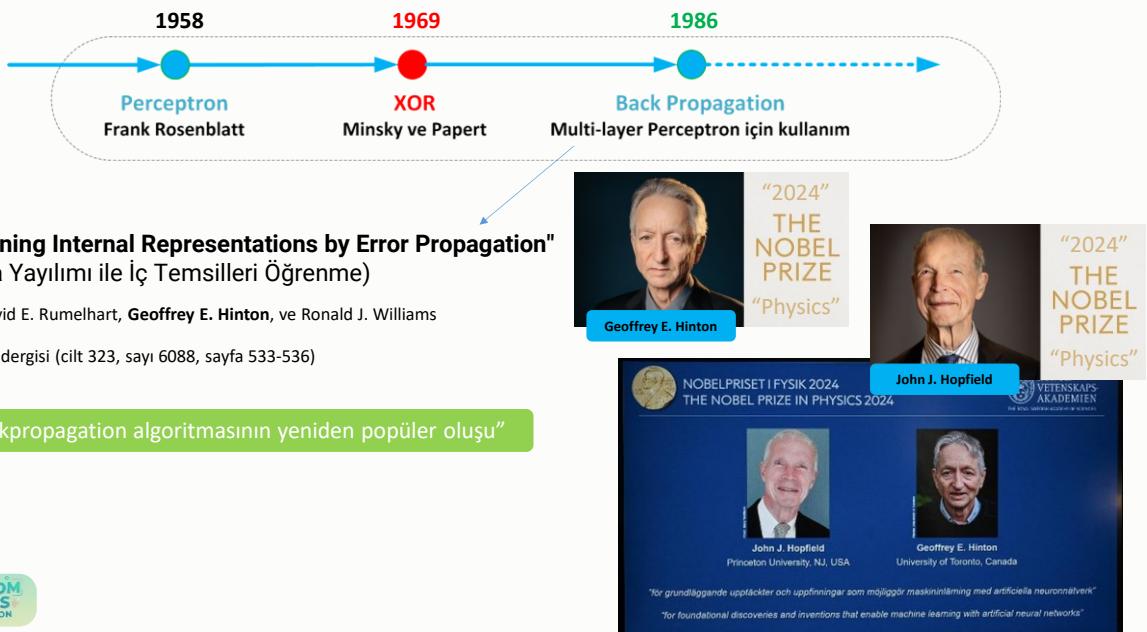
- Giriş katmanı: 2 düğüm
- 4 Gizil (hidden) katman, her birinde 10 düğüm
- Çıkış katmanı: 1 düğüm

Aktivasyon fonksiyonu: sigmoid

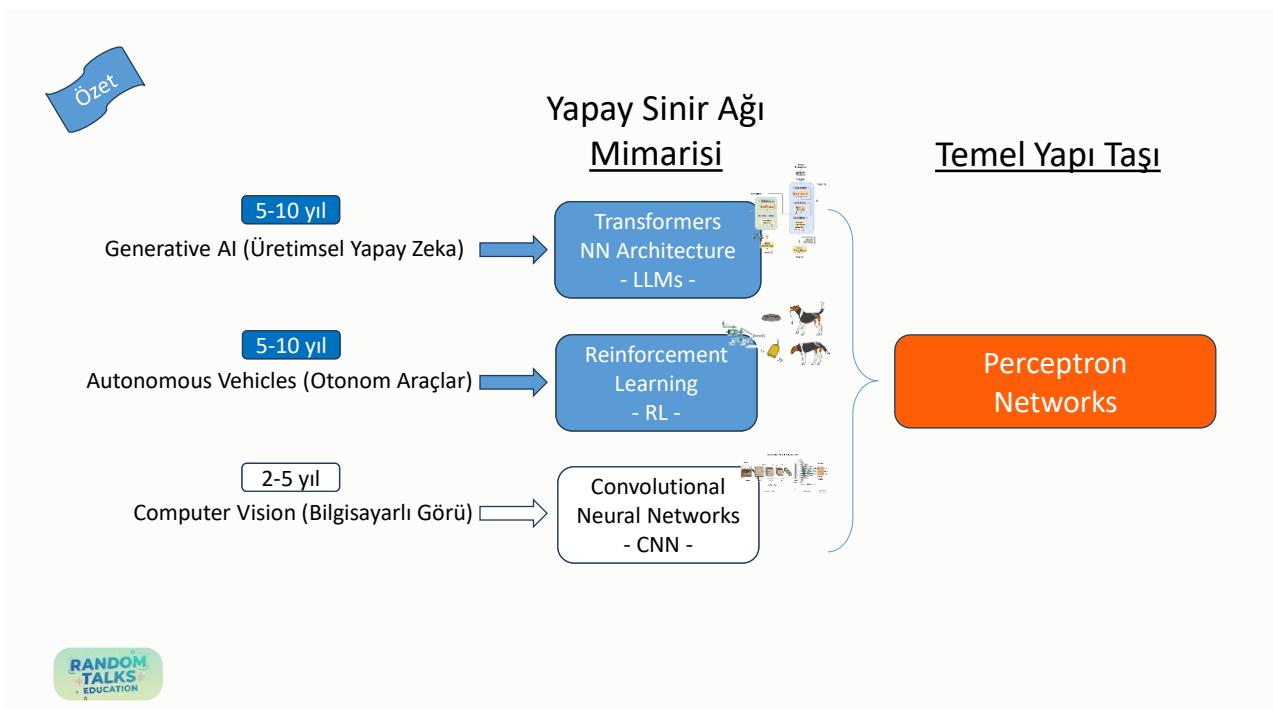
Parametre sayısı : 371



114



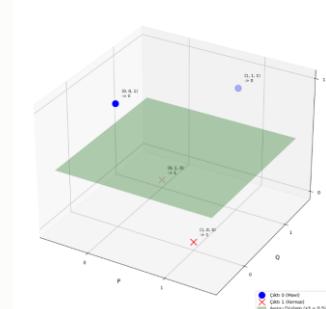
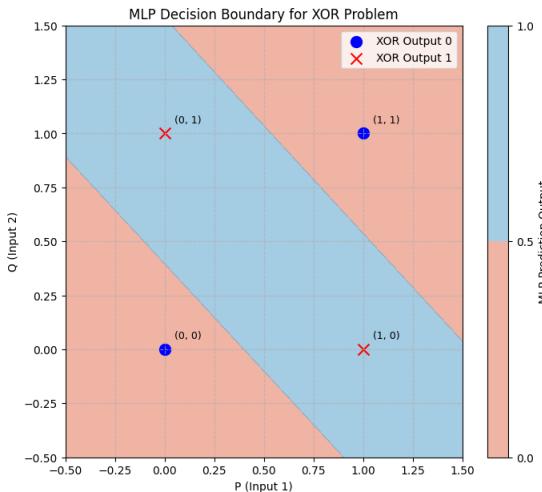
115



116

Son Söz

XOR Probleminin MLP ile Çözümü

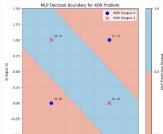


117

Son Söz

SVM Kernel Trick

(3. boyut ile noktaları ayır)



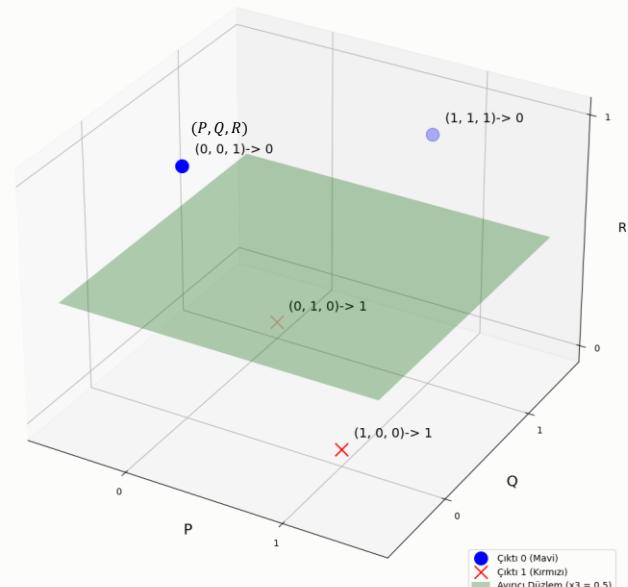
P	Q	P XOR Q
1	1	0
1	0	1
0	1	1
0	0	0

3. boyut

P	Q	R = P \Leftrightarrow Q
1	1	1
1	0	0
0	1	0
0	0	1

$$R = P \Leftrightarrow Q = \neg(P \text{ XOR } Q)$$

XOR Probleminin Doğrusal Model ile Çözümü

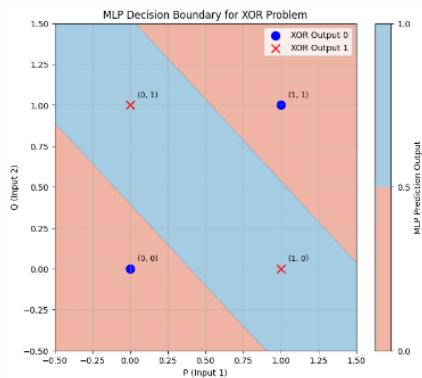


118

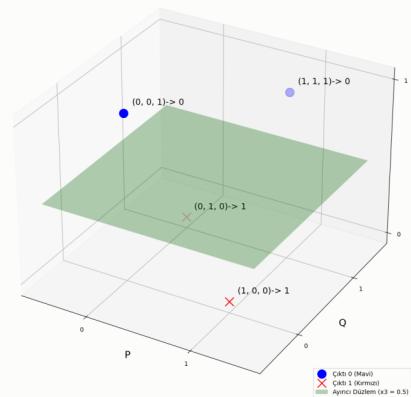


Son Söz

Zekanın Yapayı



Zekanın Doğalı



119

Random Talks Education
- Seminar Series -



<https://github.com/RandomTalksEducation/>



www.youtube.com/@RandomTalksEducation

Teşekkürler
**Yapay Sinir Ağlarının
Sıfırdan Keşfi**

Prof. Dr. Bekir Taner Dinçer

Muğla Sıtkı Koçman Üniversitesi
Mühendislik Fakültesi
Bilgisayar Mühendisliği Bölümü



120