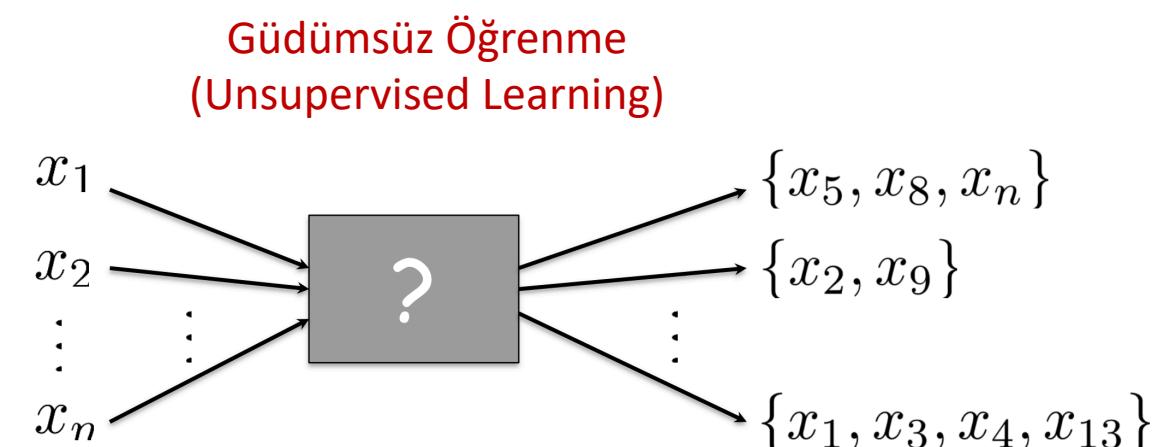
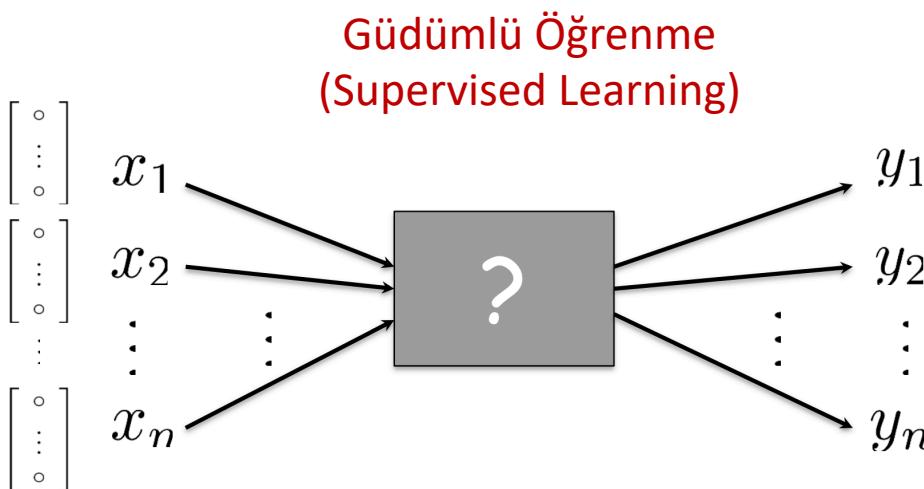
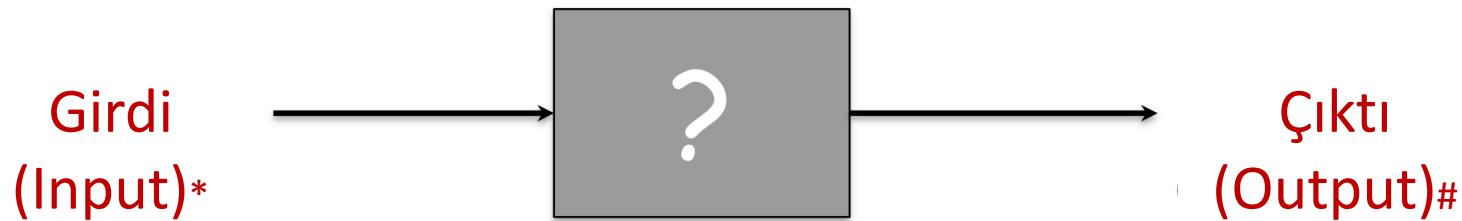




Makine Öğrenmesine Giriş

Hesaplamalı Matematik I - Özgür Martin

Makine Öğrenmesi



* Bağımsız değişken (independent variable); kestirici (predictor)

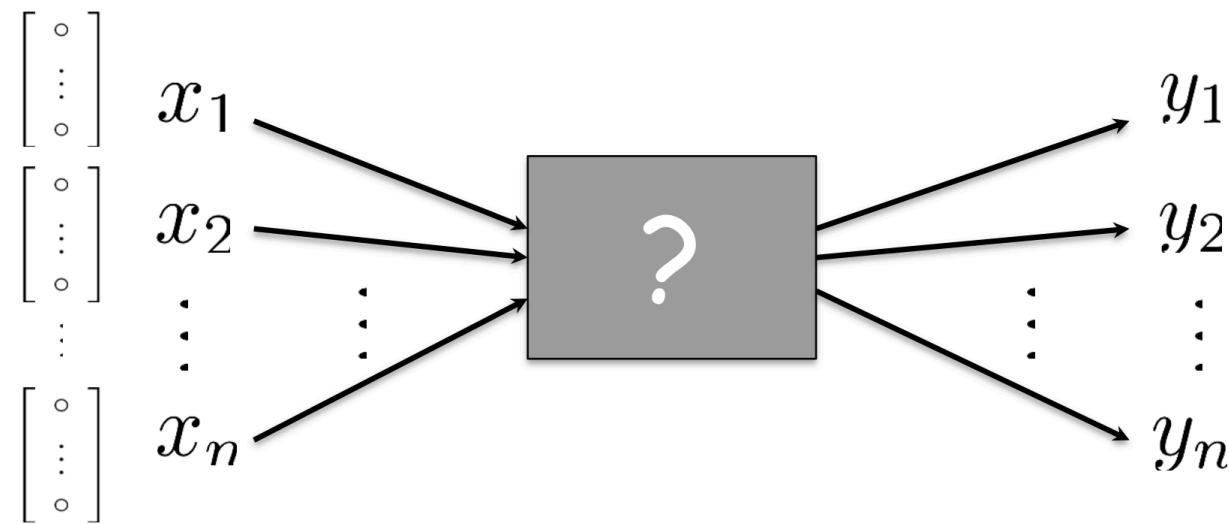
Bağımlı değişken (dependent variable); hedef değer (target value)

• İlker Birbil, IMO2020 ders notları. ([www](#))

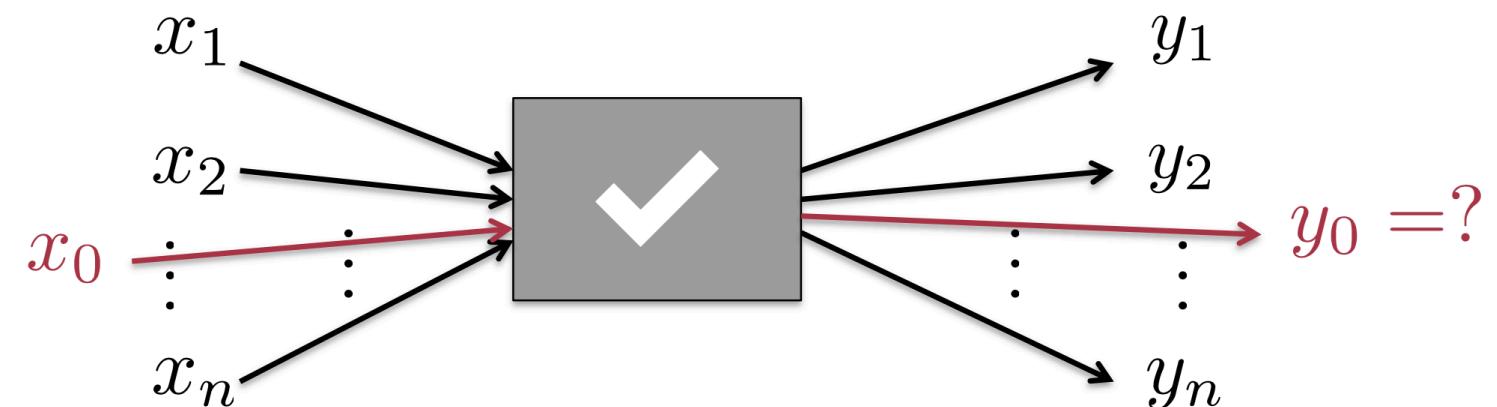
• “An Introduction to Statistical Learning – with Applications in R,”
G. James, D. Witten, T. Hastie, R. Tibshirani. 7th Ed., Springer, New York, 2013. ([www](#))

Güdümlü Öğrenme (Supervised Learning)

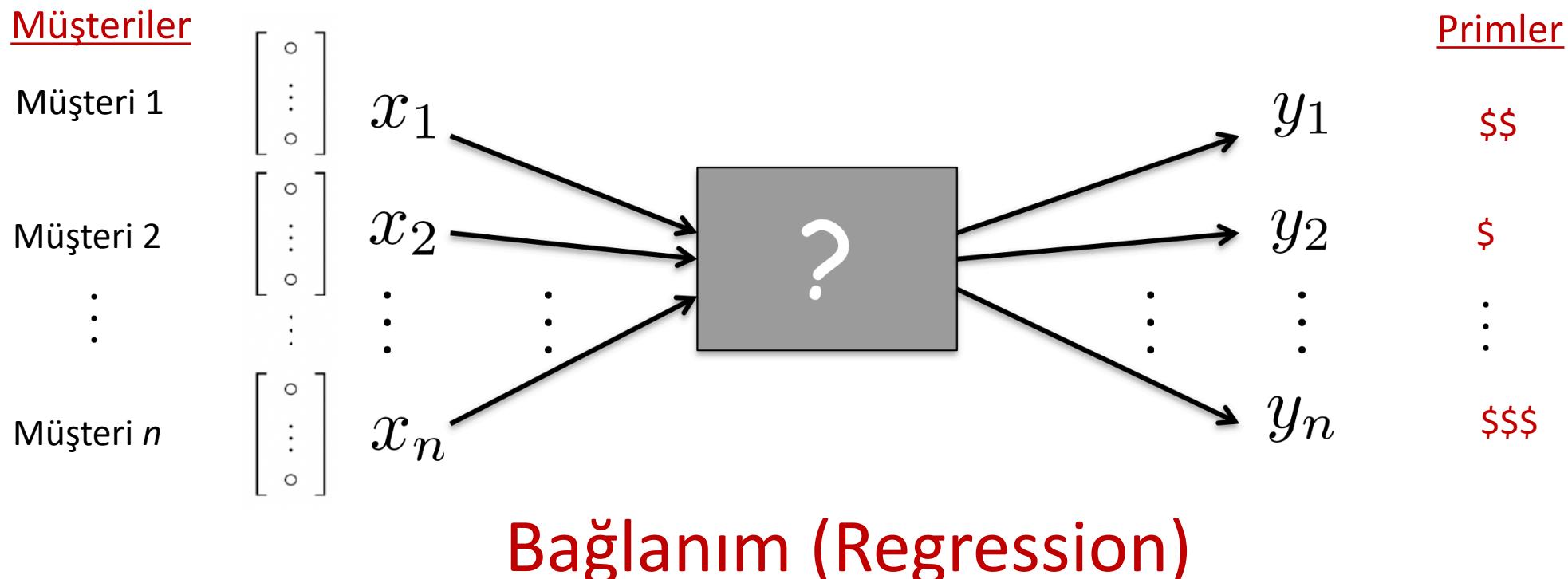
Eğitim verisi
 $\{(x_i, y_i): i = 1, \dots, n\}$



Test verisi
 (x_0, y_0)



Güdümlü Öğrenme (Supervised Learning)



Eğitim verisi
 $\{(x_i, y_i): i = 1, \dots, n\}$

Güdümlü Öğrenme (Supervised Learning)

Müşteriler

Müşteri 1

$$\begin{bmatrix} \circ \\ \vdots \\ \circ \end{bmatrix}$$

x_1

Müşteri 2

$$\begin{bmatrix} \circ \\ \vdots \\ \circ \end{bmatrix}$$

x_2

\vdots

Müşteri n

$$\begin{bmatrix} \circ \\ \vdots \\ \circ \end{bmatrix}$$

x_n

Sigorta talepleri

y_1

Kabul

y_2

Red

y_n

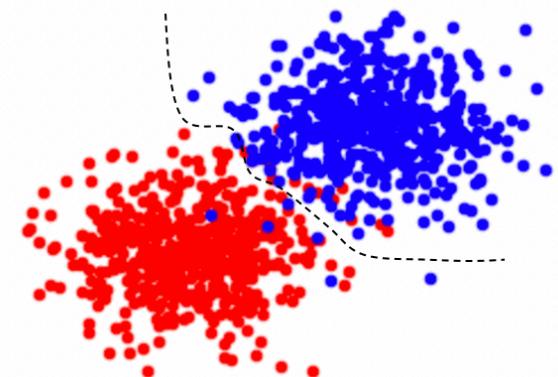
Kabul



Sınıflandırma (Classification)

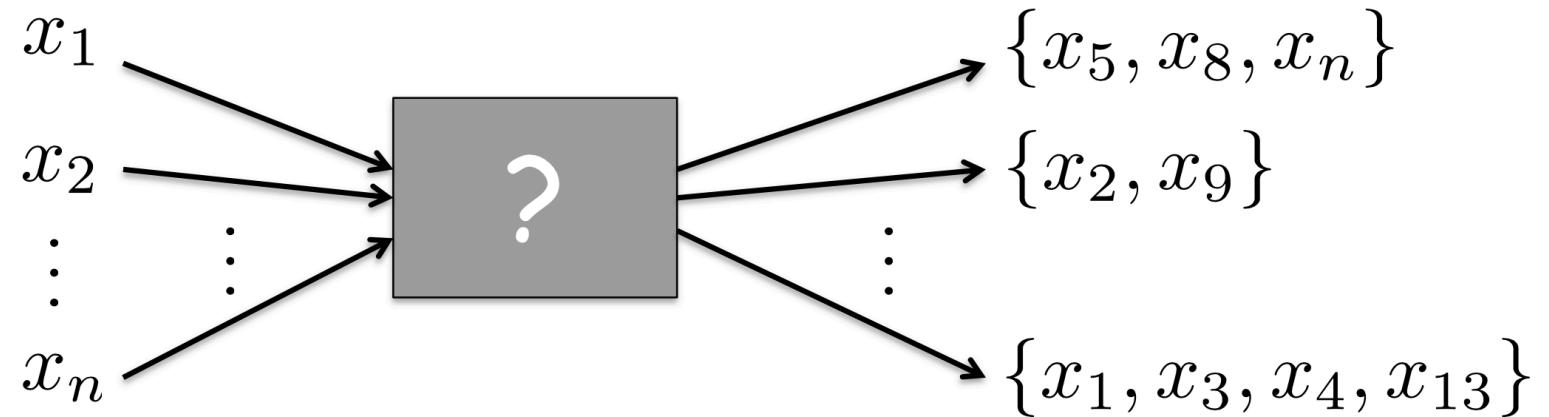
Eğitim verisi

$$\{(x_i, y_i): i = 1, \dots, n\}$$

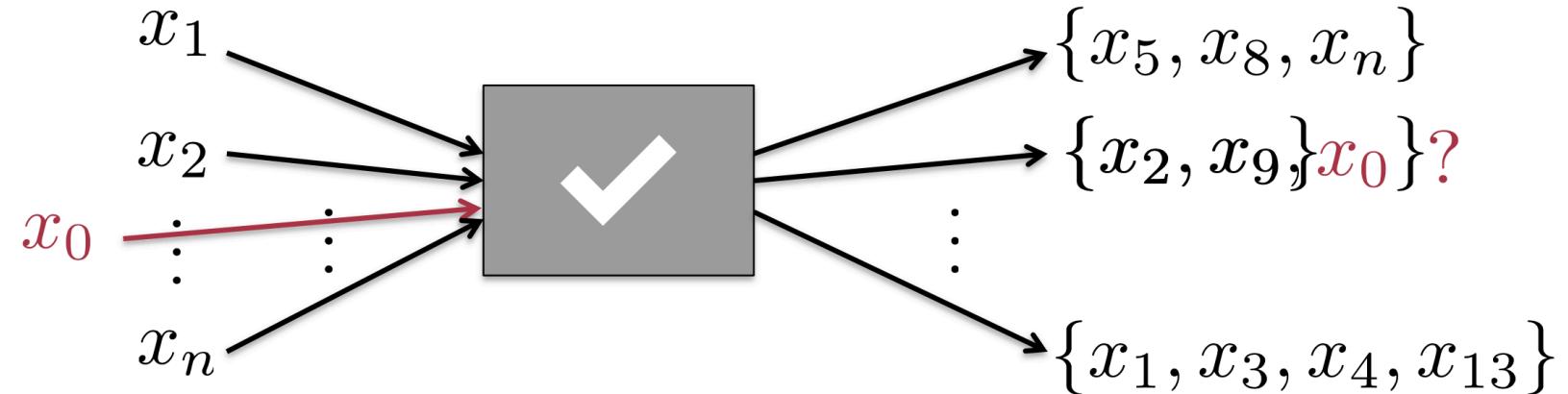


Güdümzsüz Öğrenme (Unsupervised Learning)

Eğitim verisi
 $\{x_i : i = 1, \dots, n\}$



Test verisi
 x_0



Güdümzsüz Öğrenme (Unsupervised Learning)

Customers

Customer 1

x_1

Customer 2

x_2

:

:

Customer n

x_n

?

Segmentasyon

$\{x_5, x_8, x_n\}$

$\{x_2, x_9\}$

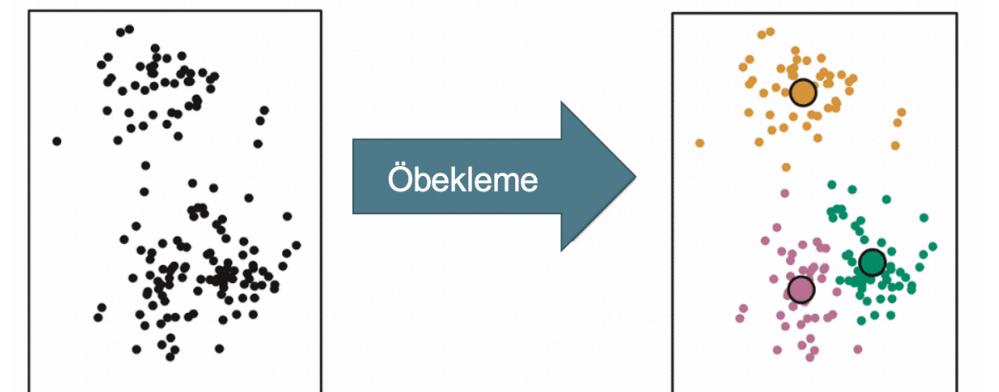
:

$\{x_1, x_3, x_4, x_{13}\}$

Öbekleme (Clustering)

Training data

$\{x_i : i = 1, \dots, n\}$



Öğrenmeden Beklentimiz

Kestirim
(Prediction)

$$x_0 = \begin{bmatrix} \circ \\ \vdots \\ \circ \end{bmatrix} \longrightarrow y_0 = ?$$

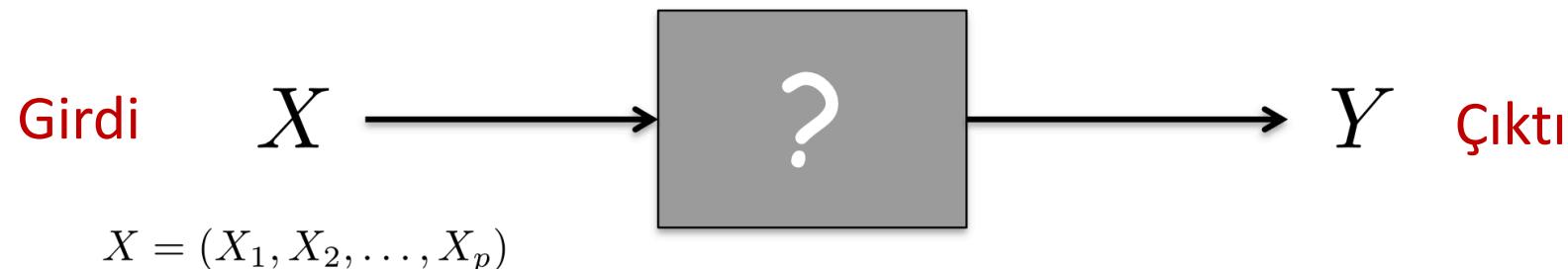
Ne?

Çıkarım
(Inference)

$$x_0 = \begin{bmatrix} \bullet \\ \vdots \\ \bullet \end{bmatrix} \xrightarrow{?} y_0$$

Nasıl?

Öğrenme Problemi

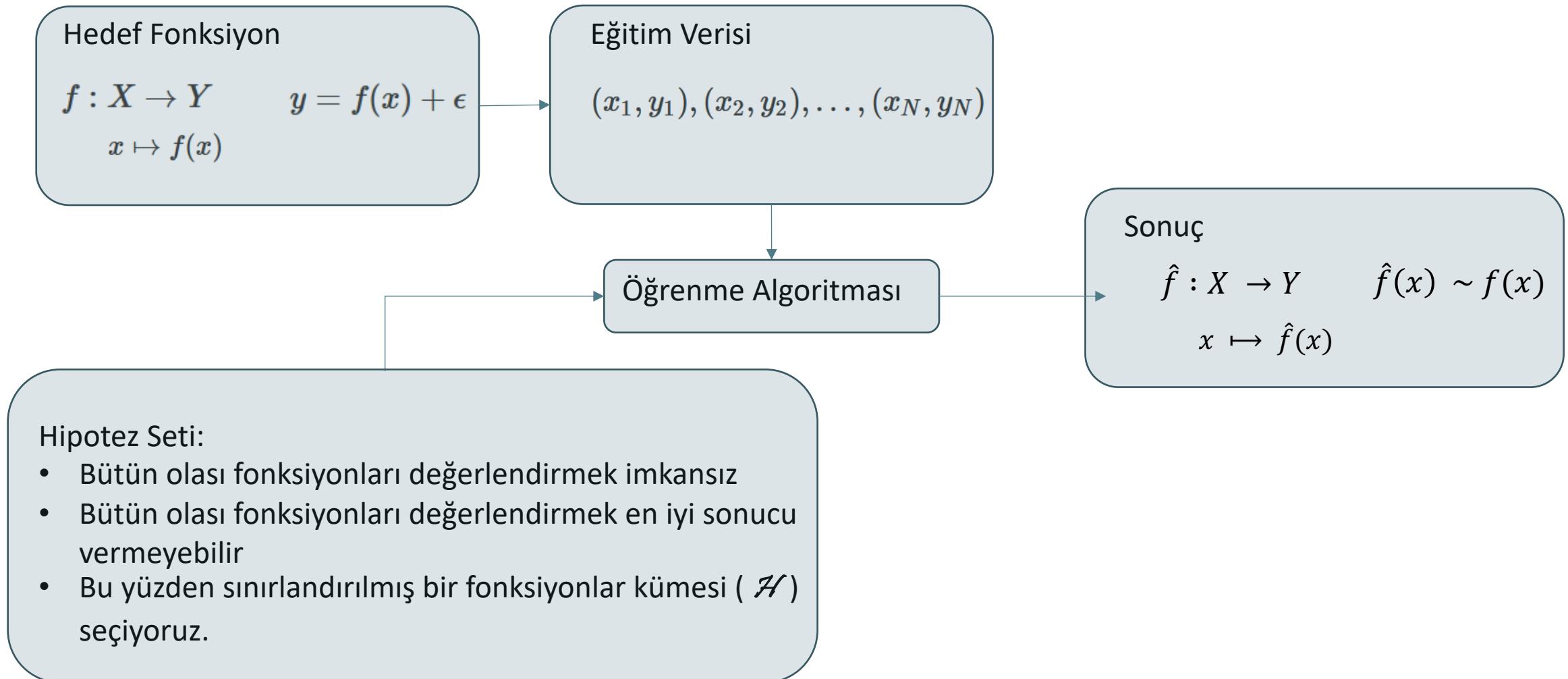


Bilinmeyen Fonksiyon

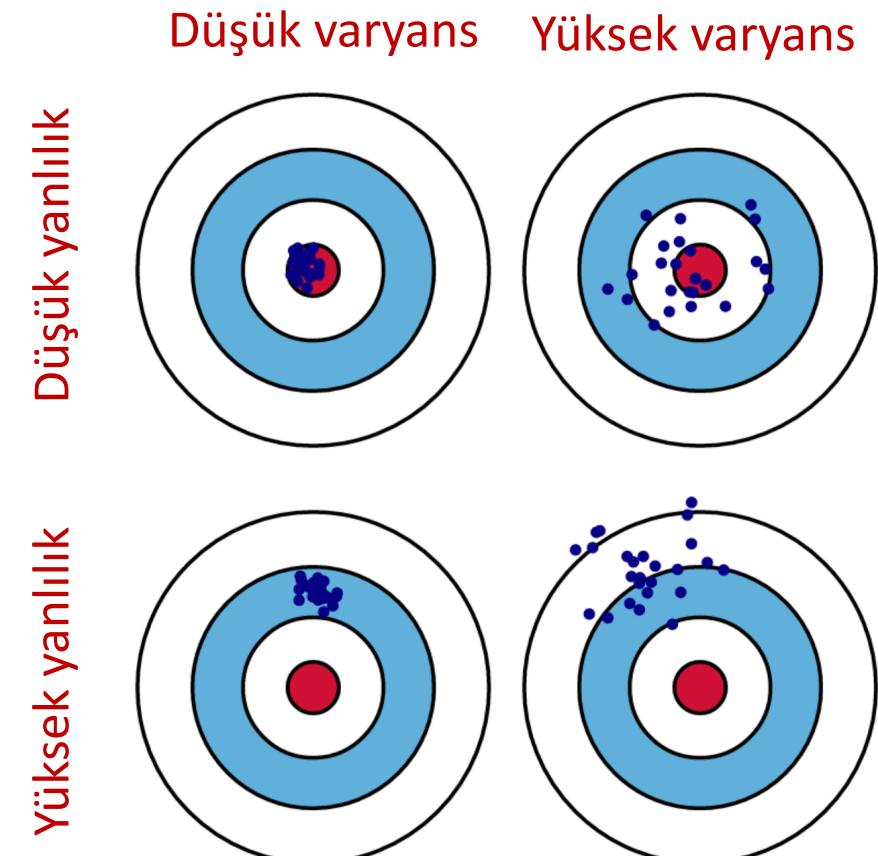
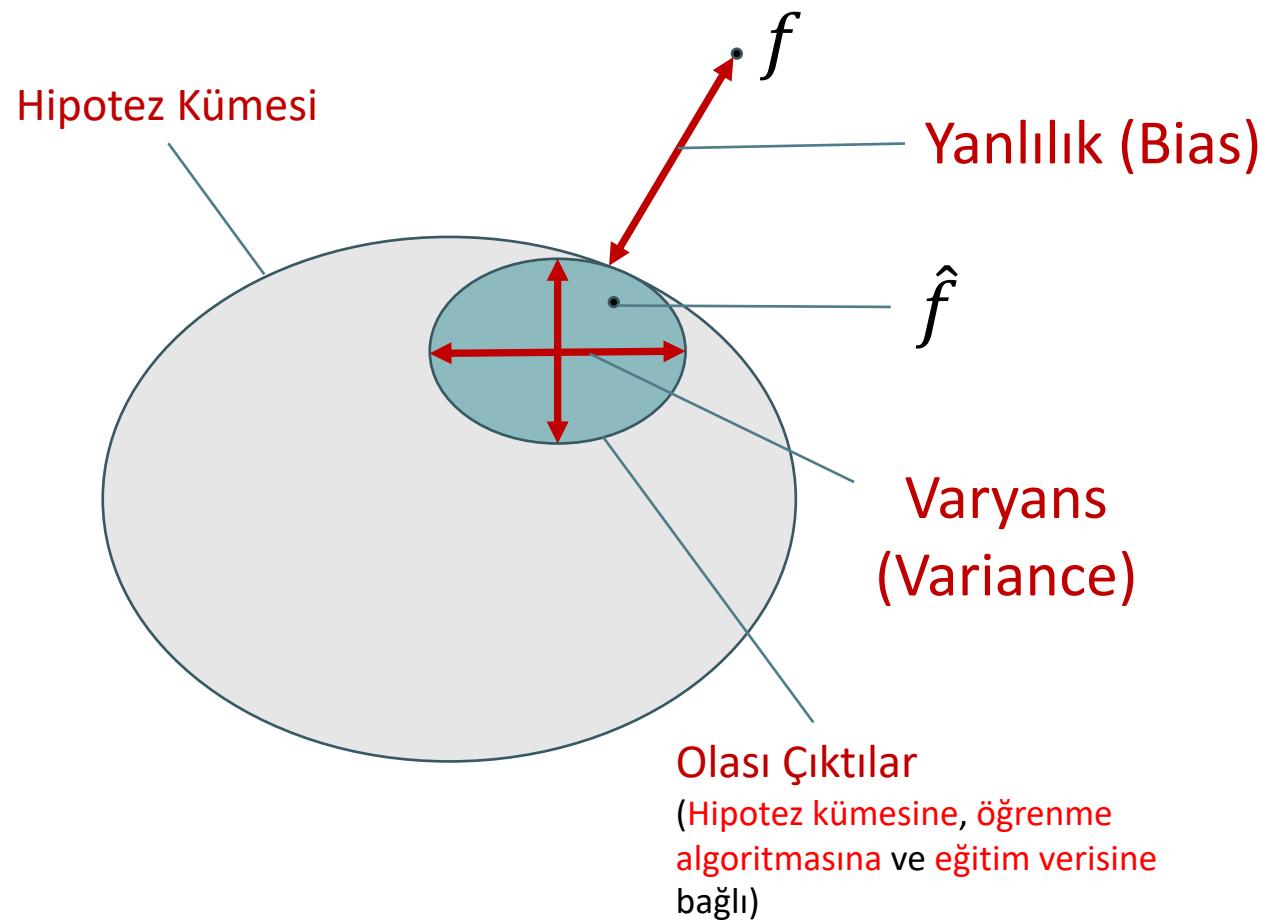
$$Y = f(X) + \epsilon \xrightarrow{\text{Yaklaşık?}} \hat{Y} = \hat{f}(X)$$

↓
Rastgele Hata Terimi
(Girdiden bağımsız)
↑

Öğrenme Problemi



Öğrenme Problemi



Modelimiz ne kadar genel olmalı?

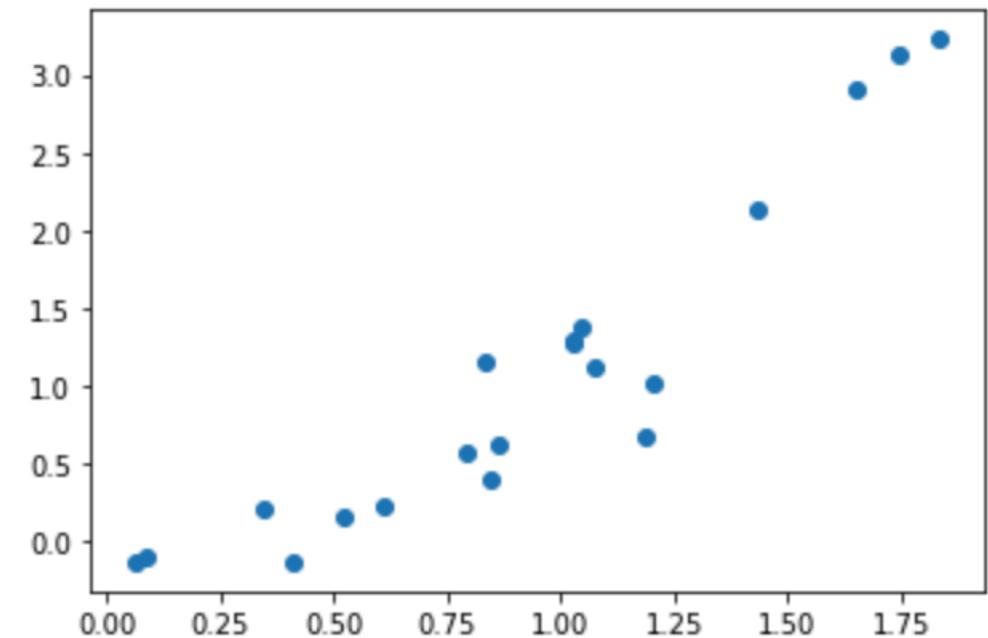
Amaç: Öyle bir fonksiyon bulalım ki

- Veriyi iyi modellesin
- Yeni veride de iyi performans sergilesin

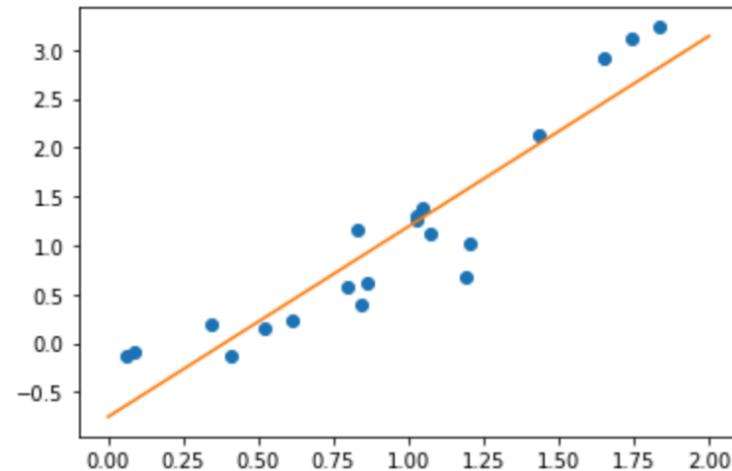
Bir polinom deneyelim

$$P(x) = c_n x^n + \cdots + c_1 x + c_0$$

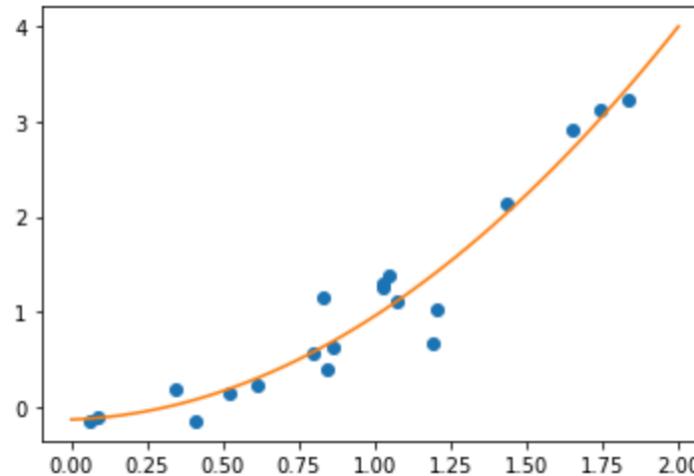
n derecesi kaç olmalı? Küçük-Büyük?



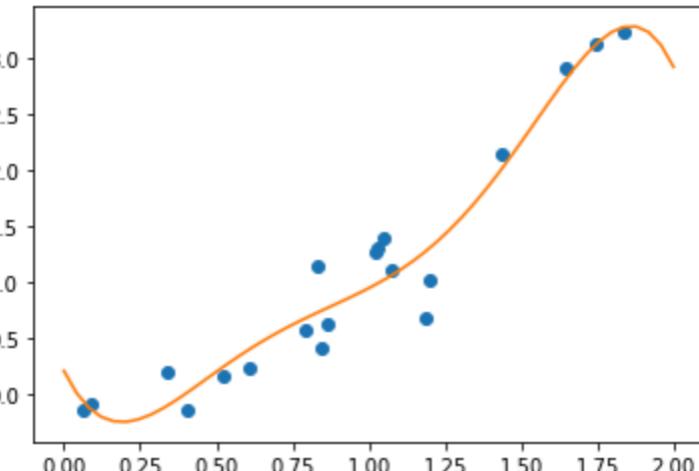
$n = 1$



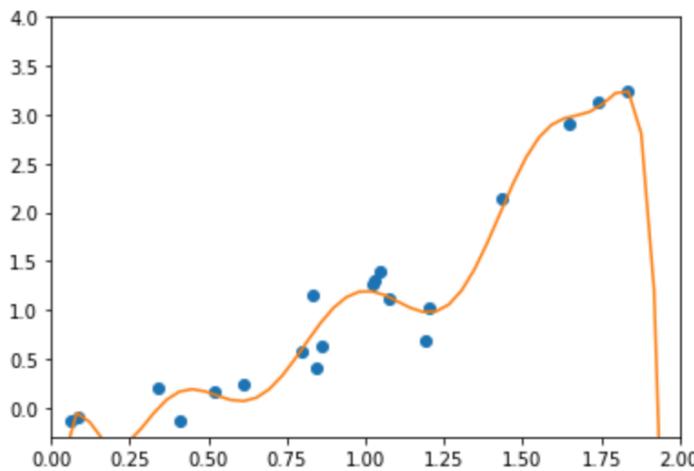
$n = 2$



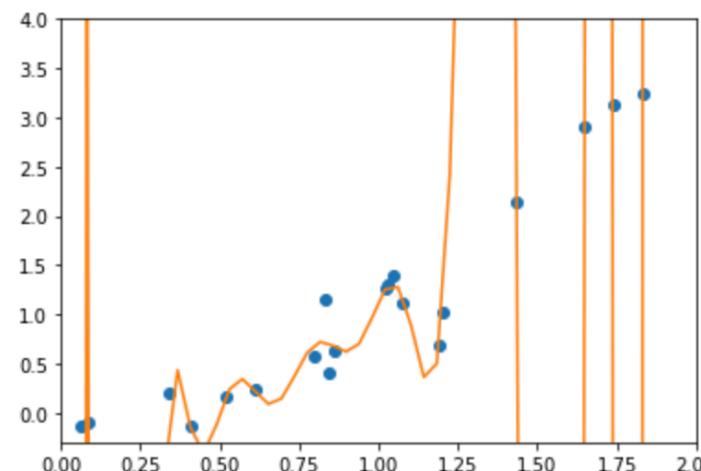
$n = 5$



$n = 10$



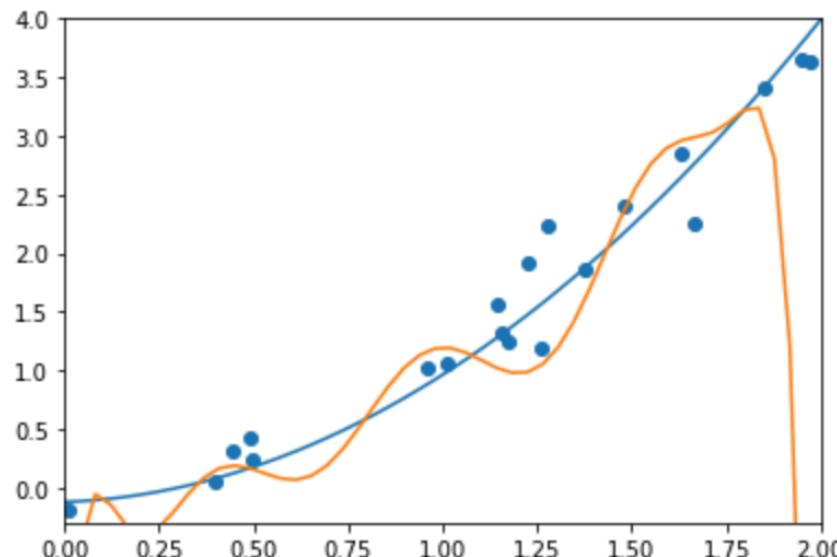
$n = 15$



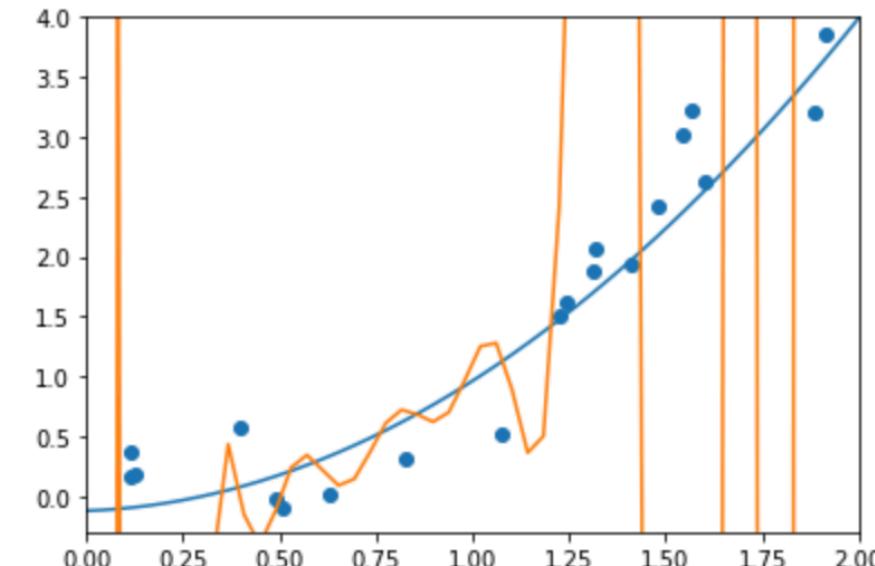
Bu veriyi şöyle bir olasılıksal fonksiyondan ürettik: $Y = X^2 + \epsilon$

Performanslarını yeni veri üzerinde deneyelim $n = 2$ vs $n = 10$ or $n = 15$

$n = 2$ vs $n = 10$



$n = 2$ vs $n = 15$



$n = 10, 15$ için aşırı öğrenme (Overfitting)!

$n \uparrow$

Varyans \uparrow

Yanlılık \downarrow