

Gözetimli Öğrenme

Veri Hazırlama

		ID	Yaş	Maaş	Şehir	Çalışma Saati	Memnuniyet	Departman	Deneyim Yılı
	0	1	25.0000	3000.000000	11	40	3	3	1
	1	2	30.0000	4500.000000	1	45	4	6	5
	2	3	35.0000	5000.000000	12	50	5	8	10
	3	4	35.1875	7000.000000	11	38	2	3	3
	4	5	40.0000	3200.0000000	1	42	3	6	7
	5	6	50.0000	8000.000000	12	55	5	8	12
	6	7	28.0000	5766.666667	11	60	1	3	2
	7	8	32.0000	6200.000000	1	48	4	6	4
	8	9	35.1875	7500.000000	12	35	2	8	6
	10	11	29.0000	4000.000000	3	39	4	12	3
	11	12	38.0000	5200.000000	0	44	3	7	8
	12	13	41.0000	5600.000000	10	52	4	10	9
	13	14	35.1875	6800.000000	5	41	5	0	5
	14	15	36.0000	7100.000000	6	53	3	11	6
	15	16	27.0000	4300.000000	8	47	2	5	2
	16	17	34.0000	3900.000000	2	49	4	1	11
	17	18	42.0000	4700.000000	9	46	5	9	14
	18	19	31.0000	5766.666667	7	50	1	4	4
ı	19	20	35.1875	8800.000000	4	65	3	2	13



Gözetimli Öğrenme

Veri Hazırlama

	ID	Yaş	Maaş	Şehir	Çalışma Saati	Memnuniyet	Departman	Deneyim Yılı
0	1	25.0000	3000.000000	11	40	3	3	1
1	2	30.0000	4500.000000	1	45	4	6	5
2	3	35.0000	5000.000000	12	50	5	8	10
3	4	35.1875	7000.000000	11	38	2	3	3
4	5	40.0000	3200.0000000	1	42	3	6	7
5	6	50.0000	8000.000000	12	55	5	8	12
6	7	28.0000	5766.666667	11	60	1	3	2
7	8	32.0000	6200.000000	1	48	4	6	4
8	9	35.1875	7500.000000	12	35	2	8	6
10	11	29.0000	4000.000000	3	39	4	12	3
11	12	38.0000	5200.000000	0	44	3	7	8
12	13	41.0000	5600.000000	10	52	4	10	9
13	14	35.1875	6800.000000	5	41	5	0	5
14	15	36.0000	7100.000000	6	53	3	11	6
15	16	27.0000	4300.000000	8	47	2	5	2
16	17	34.0000	3900.000000	2	49	4	1	11
17	18	42.0000	4700.000000	9	46	5	9	14
18	19	31.0000	5766.666667	7	50	1	4	4
19	20	35.1875	8800.000000	4	65	3	2	13
	·			•				

Öznitelik Mühendisliği

- Öznitelik Seçimi (Feature Selection)
- Öznitelik Oluşturma (Feature Extraction)



Gözetimli Öğrenme

Veri Hazırlama

V/2.5						
Yaş	Maaş	Şehir	Çalışma Saati	Memnuniyet	Departman	Deneyim Yılı
25.0000	3000.000000	11	40		3	1
30.0000	4500.000000	1	45	4	6	5
35.0000	5000.000000	12	50	5	8	10
35.1875	7000.000000	11	38	2	3	3
40.0000	3200.0000000	1	42		6	7
50.0000	8000.000000	12	55	5	8	12
28.0000	5766.666667	11	60		3	2
32.0000	6200.000000	1	48	4	6	4
35.1875	7500.000000	12	35	2	8	6
29.0000	4000.000000	3	39	4	12	3
38.0000	5200.000000	0	44		7	8
41.0000	5600.000000	10	52		10	9
35.1875	6800.000000	5	41	5	0	5
36.0000	7100.000000	6	53		11	6
27.0000	4300.000000	8	47	2	5	2
34.0000	3900.000000	2	49	4	1	11
42.0000	4700.000000	9	46	5	9	14
31.0000	5766.666667	7	50		4	4
35 . 1875	8800.000000	4	65		2	13
	30.0000 35.0000 35.1875 40.0000 28.0000 32.0000 35.1875 29.0000 38.0000 41.0000 35.1875 36.0000 27.0000 34.0000 42.0000 31.0000	30.0000 4500.000000 35.0000 5000.000000 35.1875 7000.000000 40.0000 3200.000000 50.0000 8000.000000 28.0000 5766.666667 32.0000 6200.000000 35.1875 7500.000000 29.0000 4000.000000 38.0000 5200.000000 41.0000 5600.000000 35.1875 6800.000000 36.0000 7100.000000 27.0000 4300.000000 34.0000 3900.000000 42.0000 4700.000000 31.0000 5766.666667	30.0000 4500.000000 1 35.0000 5000.000000 12 35.1875 7000.000000 11 40.0000 3200.000000 1 50.0000 8000.000000 12 28.0000 5766.666667 11 32.0000 6200.000000 12 29.0000 4000.000000 3 38.0000 5200.000000 0 41.0000 5600.000000 10 35.1875 6800.000000 5 36.0000 7100.000000 6 27.0000 4300.000000 8 34.0000 3900.000000 2 42.0000 4700.000000 9 31.0000 5766.666667 7	30.0000 4500.000000 1 45 35.0000 5000.000000 12 50 35.1875 7000.000000 11 38 40.0000 3200.000000 1 42 50.0000 8000.000000 1 42 28.0000 5766.666667 11 60 32.0000 6200.000000 1 48 35.1875 7500.000000 12 35 29.0000 4000.000000 3 39 38.0000 5200.000000 0 44 41.0000 5600.000000 10 52 35.1875 6800.000000 5 41 36.0000 7100.000000 6 53 27.0000 4300.000000 8 47 34.0000 3900.000000 2 49 42.0000 4700.000000 9 46 31.0000 5766.6666667 7 50	30.0000 4500.000000 1 45 4 35.0000 5000.000000 12 50 5 35.1875 7000.000000 11 38 2 40.0000 3200.000000 1 42 3 50.0000 8000.000000 12 55 5 28.0000 5766.6666667 11 60 1 32.0000 6200.000000 1 48 4 35.1875 7500.000000 12 35 2 29.0000 4000.000000 3 39 4 38.0000 5200.000000 0 44 3 41.0000 5600.000000 10 52 4 35.1875 6800.000000 5 41 5 36.0000 7100.000000 6 53 3 27.0000 4300.000000 8 47 2 34.0000 3900.000000 2 49 4 42.0000 4700.000000 9 46 5 31.0000 5766.6666667 7 50 </td <td>30.0000 4500.000000 1 45 4 6 35.0000 5000.000000 12 50 5 8 35.1875 7000.000000 11 38 2 3 40.0000 3200.000000 1 42 3 6 50.0000 8000.000000 12 55 5 8 28.0000 5766.666667 11 60 1 3 32.0000 6200.000000 1 48 4 6 35.1875 7500.000000 12 35 2 8 29.0000 4000.000000 3 39 4 12 38.0000 5200.000000 0 44 3 7 41.0000 5600.000000 10 52 4 10 35.1875 6800.000000 5 41 5 0 36.0000 7100.000000 6 53 3 11 27.0000 4300.000000 8 47 2 5 34.0000 3900.000000 2 49</td>	30.0000 4500.000000 1 45 4 6 35.0000 5000.000000 12 50 5 8 35.1875 7000.000000 11 38 2 3 40.0000 3200.000000 1 42 3 6 50.0000 8000.000000 12 55 5 8 28.0000 5766.666667 11 60 1 3 32.0000 6200.000000 1 48 4 6 35.1875 7500.000000 12 35 2 8 29.0000 4000.000000 3 39 4 12 38.0000 5200.000000 0 44 3 7 41.0000 5600.000000 10 52 4 10 35.1875 6800.000000 5 41 5 0 36.0000 7100.000000 6 53 3 11 27.0000 4300.000000 8 47 2 5 34.0000 3900.000000 2 49

Öznitelik Mühendisliği

- Öznitelik Seçimi (Feature Selection)
- Öznitelik Oluşturma (Feature Extraction)



Gözetimli Öğrenme

Veri Hazırlama

	ID	Yaş	Maaş	Şehir	Çalışma Saati	Memnuniyet	Departman	Deneyim Yılı
0		25.0000	3000.000000	11	40		3	1
1	2	30.0000	4500.000000	1	45	4	6	5
2		35.0000	5000.000000	12	50	5	8	10
3		35.1875	7000.000000	11	38	2	3	3
4	5	40.0000	3200.000000	1	42		6	7
5	6	50.0000	8000.000000	12	55	5	8	12
6		28.0000	5766.666667	11	60		3	2
7		32.0000	6200.000000	1	48	4	6	4
8	9	35.1875	7500.000000	12	35	2	8	6
10	11	29.0000	4000.000000	3	39	4	12	3
11	12	38.0000	5200.000000	0	44		7	8
12	13	41.0000	5600.000000	10	52		10	9
13	14	35.1875	6800.000000	5	41	5	0	5
14	15	36.0000	7100.000000	6	53		11	6
15	16	27.0000	4300.000000	8	47	2	5	2
16	17	34.0000	3900.000000	2	49	4	1	11
17	18	42.0000	4700.000000	9	46	5	9	14
18	19	31.0000	5766.666667	7	50		4	4
19	20	35.1875	8800.000000	4	65		2	13

Öznitelik Mühendisliği

- Öznitelik Seçimi (Feature Selection)
- Öznitelik Oluşturma (Feature Extraction)



Gözetimli Öğrenme

Veri Hazırlama

011 1 1A2		14		
Yaş	Şehir	Çalışma Saati	Departman	Deneyim Yılı
25.0000	11	40	3	1
30.0000	1	45	6	5
35.0000	12	50	8	10
35.1875	11	38	3	3
40.0000	1	42	6	7
50.0000	12	55	8	12
28.0000	11	60	3	2
32.0000	1	48	6	4
35.1875	12	35	8	6
29.0000	3	39	12	3
38.0000	0	44	7	8
41.0000	10	52	10	9
35.1875	5	41	0	5
36.0000	6	53	11	6
27.0000	8	47	5	2
34.0000	2	49	1	11
42.0000	9	46	9	14
31.0000	7	50	4	4
35.1875	4	65	2	13



Gözetimli Öğrenme

Veri Hazırlama

Yaş Şehir Çalışma Saati Departman Deneyim Y 25.0000 11 40 3 30.0000 1 45 6	1 5
30.0000 1 45 6	
	,
35 . 0000 12 50 8	10
35.1875 11 38 3	3
40.0000 1 42 6	э 7
50.0000 12 55 8	12
28.0000 11 60 3	2
32.0000 1 48 6	4
35.1875 12 35 8	6
29.0000 3 39 12	3
38.0000 0 44 7	8
41.0000 10 52 10	9
35.1875 5 41 0	5
36.0000 6 53 11	6
27.0000 8 47 5	2
34.0000 2 49 1	11
42.0000 9 46 9	14
31.0000 7 50 4	4
35.1875 4 65 2	13

Veri Normalizasyonu ve Standardizasyonu

- Min-Max Normalization
- Z-Skor Normalizasyonu
- · Robust skalalama
- Kısa süreli ortalama ve varyans normalizasyonu
- Kısa süreli kepstral ve ölçek normalizasyonu



Gözetimli Öğrenme

Veri Hazırlama

Veri Normalizasyonu ve Standardizasyonu

- Min-Max Normalization
- Z-Skor Normalizasyonu
- Robust skalalama
- Kısa süreli ortalama ve varyans normalizasyonu
- Kısa süreli kepstral ve ölçek normalizasyonu



Gözetimli Öğrenme

Veri Hazırlama

Veri Normalizasyonu ve Standardizasyonu

- Min-Max Normalization
- Z-Skor Normalizasyonu
- Robust skalalama
- Kısa süreli ortalama ve varyans normalizasyonu
- Kısa süreli kepstral ve ölçek normalizasyonu



Gözetimli Öğrenme

Veri Hazırlama

Veri Normalizasyonu ve Standardizasyonu

Min-Max Normalization

$$X_{
m normalized} = rac{X - X_{
m min}}{X_{
m max} - X_{
m min}}$$

Avantajları: Kolay anlaşılır ve ölçeklendirilmiş veriler sabit bir aralığa gelir.

Dezavantajları: Aykırı değerlere duyarlıdır.



Gözetimli Öğrenme

Veri Hazırlama

Veri Normalizasyonu ve Standardizasyonu

Min-Max Normalization

$$X_{
m normalized} = rac{X - X_{
m min}}{X_{
m max} - X_{
m min}}$$

$$X = [1, 2, 8, 4, 5, 10]$$



Gözetimli Öğrenme

Veri Hazırlama

Veri Normalizasyonu ve Standardizasyonu

Min-Max Normalization

$$X_{
m normalized} = rac{X - X_{
m min}}{X_{
m max} - X_{
m min}}$$

$$X = [1, 2, 8, 4, 5, 11]$$

Örnek X = 8 olsun

Xmin = 11

 $X_{max} = 1$



Gözetimli Öğrenme

Veri Hazırlama

Veri Normalizasyonu ve Standardizasyonu

Min-Max Normalization

$$X = [1, 2, 8, 4, 5, 11]$$

Örnek X = 8 olsun

Xmin = 11

 $X_{max} = 1$

$$X_{
m normalized} = rac{X - X_{
m min}}{X_{
m max} - X_{
m min}}$$

$$X_{
m normalized} = rac{8-1}{11-1} = rac{7}{10} = 0.7$$



Gözetimli Öğrenme

Veri Hazırlama

Veri Normalizasyonu ve Standardizasyonu

Min-Max Normalization

$$X = [1, 2, 8, 4, 5, 11]$$

Örnek X = 8 olsun

Xmin = 11

 $X_{max} = 1$

$$X_{
m normalized} = rac{X - X_{
m min}}{X_{
m max} - X_{
m min}}$$

$$X_{
m normalized} = rac{8-1}{11-1} = rac{7}{10} = 0.7$$



Gözetimli Öğrenme

Veri Hazırlama

Veri Normalizasyonu ve Standardizasyonu

Min-Max Normalization

$$X = [1, 2, 8, 4, 5, 11, 700]$$

Örnek X = 8 olsun

$$Xmin = 700$$

$$X_{max} = 1$$

$$X_{
m normalized} = rac{X - X_{
m min}}{X_{
m max} - X_{
m min}}$$

$$egin{aligned} X_{ ext{normalized}} &= rac{8-1}{700-1} \ &= rac{7}{699} \ &pprox 0.0100 \end{aligned}$$



Gözetimli Öğrenme

Veri Hazırlama

Veri Normalizasyonu ve Standardizasyonu

Min-Max Normalization

$$X_{
m normalized} = rac{X - X_{
m min}}{X_{
m max} - X_{
m min}}$$

X (Orijinal)	X (Normalize Edilmiş)
1	0.0000
2	0.0014
8	0.0100
4	0.0043
5	0.0057
11	0.0143
700	1.0000

★ Gözlem: 700 gibi büyük bir değer eklendiği için diğer tüm değerler sıfıra çok yakın hale geldi. [-]



Gözetimli Öğrenme

Veri Hazırlama

Veri Normalizasyonu ve Standardizasyonu

Z-skor Normalizasyonu

$$X_{ ext{standardized}} = rac{X - \mu}{\sigma}$$

Burada μ veri kümesinin ortalaması, σ ise standart sapmasıdır.

- Avantajları: Verileri normal dağılıma yaklaştırır ve aykırı değerlere karşı daha dayanıklıdır.
- Dezavantajları: Verilerin dağılımı normal değilse etkisi azalabilir.



Gözetimli Öğrenme

Veri Hazırlama

Veri Normalizasyonu ve Standardizasyonu

Robust Skalalama

$$X_{ ext{robust}} = rac{X - ext{Medyan}}{ ext{IQR}}$$

- Avantajları: Aykırı değerlere karşı dayanıklıdır.
- Dezavantajları: Eğer veri çok küçük bir aralıkta yoğunlaşmışsa etkisi az olabilir.



Gözetimli Öğrenme

Veri hazırlama

