



# ISE 302 –Veri Madenciliđi

DR. ÖĐR. ÜYESİ ESİN AYŞE ZAIMOĐLU



esinzaimoglu@sakarya.edu.tr

# Normalizasyon

---

Farklı ölçekleme sisteminde bulunan verilerin birbiri ile karşılaştırılabilmesi için yapılır.

Amaç, matematiksel fonksiyonlar kullanarak, farklı sistemlerde bulunan verileri, ortak bir sisteme taşımak ve karşılaştırılabilir hale getirmektir.

# Asgari – Azami Normalleştirme (Min-Max Normalisation)

---

- ❑ Bu yöntemde, bir grup verinin içerisindeki en büyük ve en küçük değerler ele alınır. Diğer bütün veriler, bu değerlere göre normalleştirilir.
- ❑ Amaç en küçük değeri 0 ve en büyük değeri 1 olacak şekilde normalleştirmek ve diğer bütün verileri bu 0-1 aralığına yaymaktır.

# Örnek

$$\frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

5, 8, 9, 11, 20, 22, 24, 25, 27, 29

$$X_{\text{normalizasyon}} = (X - X_{\min}) / (X_{\max} - X_{\min})$$

- X değeri değişkendeki minimum değer olduğunda, pay 0 olacaktır ve dolayısıyla X = 0 olacaktır.
- Öte yandan, X'in değeri değişkendeki maksimum değer olduğunda, pay paydaya eşittir olacağından  $X = 1$  sonucu çıkacaktır.
- Dolayısıyla X'in değeri minimum ve maksimum değer arasındaysa, X'in değeri 0 ile 1 arasında anlamına gelir.

# Örnek

$$\frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

	5	0
	8	0,125
	9	0,166666667
	11	0,25
	20	0,625
	22	0,708333333
	24	0,791666667
	25	0,833333333
	29	1

10	0
15	0,1
16	0,12
17	0,14
50	0,8
51	0,82
55	0,9
56	0,92
60	1

# Z-Skor Normalizasyonu (Standartlaştırma)

Z-Skor Normalizasyonu (Standartlaştırma): Veri setinin her bir özelliğinin ortalaması 0 ve standart sapması 1 olacak şekilde dönüştürülmesini sağlar. Formül şu şekildedir:

$$z = \frac{x_i - \mu}{\sigma}$$

Handwritten annotations in blue ink:

- An arrow points from  $\mu$  to the word "ortalama" (mean).
- An arrow points from  $\sigma$  to the words "standart Sapma" (standard deviation).

# Z-Skor Normalizasyonu (Standartlaştırma)

İlk olarak, her bir özelliğin (notun) ortalamasını ve standart sapmasını hesaplayacağız.

Matematik notları için:

- Ortalama ( $\mu$ ) =  $(75 + 85 + 60 + 90 + 70 + 88 + 65 + 82 + 78 + 95) / 10 = 79.8$
- Standart Sapma ( $\sigma$ ) =  $\sqrt{((75 - 79.8)^2 + (85 - 79.8)^2 + \dots + (95 - 79.8)^2) / 10} \approx 10.86$

Fizik notları için de benzer şekilde hesaplanır:

- Ortalama ( $\mu$ ) =  $(82 + 88 + 70 + 92 + 78 + 85 + 72 + 80 + 75 + 90) / 10 = 81.2$
- Standart Sapma ( $\sigma$ )  $\approx 6.9$

Öğrenci	Matematik Notu	Fizik Notu
1	75	82
2	85	88
3	60	70
4	90	92
5	70	78
6	88	85
7	65	72
8	82	80
9	78	75
10	95	90

# Z-Skor Normalizasyonu (Standartlaştırma)

---

Şimdi her bir notu standartlaştırarak Z-Skorlarını hesaplayabiliriz:

Matematik notları için:

- Z-Skor ( $X_{std}$ ) = (Not - Ortalama) / Standart Sapma

Örneğin, ilk öğrencinin matematik notunun Z-Skoru:  $X_{std} = (75 - 79.8) / 10.86 \approx -0.44$

Örneğin, ilk öğrencinin fizik notunun Z-Skoru:

$$X_{std} = (82 - 81.2) / 6.9 \approx 0.12$$

Matematik Z-Skorları:

[-0.44, 0.48, -1.83, 0.98, -0.9, 0.75, -1.37, 0.2, -0.17, 1.4]

Fizik Z-Skorları:

[0.12, 0.98, -1.62, 1.57, -0.47, 0.54, -1.33, -0.17, -0.9, 1.27]