

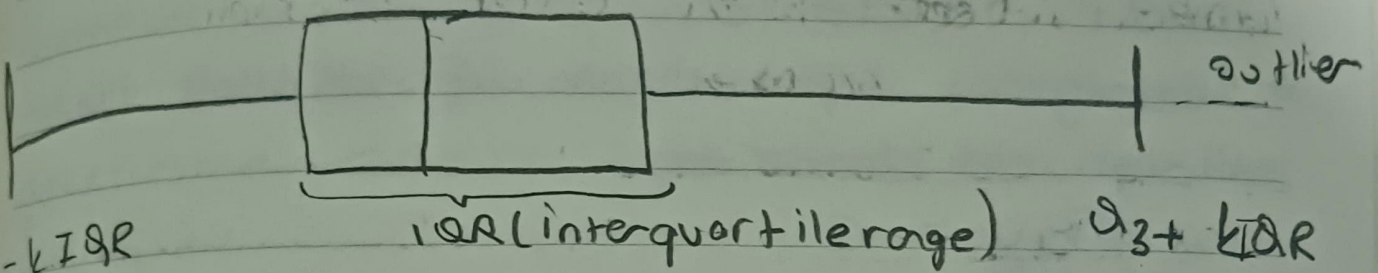
Aykırı Değer Tespit Etme Yöntemleri

outlier

• Standart sapma Yöntemi

• Box Plot (Kutu Grafiği) Yöntemi

Q_1 Q_2 Q_3
%25 %50 %75



Q_1 - LQR

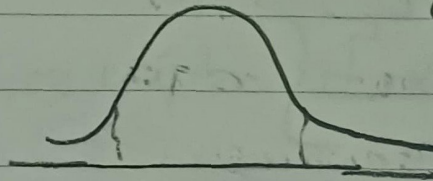
lower extreme

IQR (Interquartilerange)

Q_3 + LQR

upper extreme

• Z-Skoru Yöntemi



standart

sapmadan 2 skoru üretiriz. Veri noktasının ortalamadan ne kadar uzak olduğunu standart sapma oranıyla ölçer.

• LOF (Local Outlier Factor) : Her veri noktasının çevresindeki yoğunluğu dikkate alır. Beklenen den daha yoğun ortamda bulunan veriler aykırıdır.

④ ML için Temel matematik.

LINEER CEBİR

• Skaler : Sabit değerler. Sayılarıyla ifade edilir.

• Vektör : Hem boyutluğu hem de yönü vardır.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 \cdot 3 + 2 \cdot 5 = 13 & 1 \cdot 4 + 2 \cdot 6 = 16 \\ 3 \cdot 3 + 4 \cdot 5 = 29 & 3 \cdot 4 + 4 \cdot 6 = 36 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad A^{-1} = \frac{1}{ad-bc} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Lineer Denklemler
 * Bilinmeyen değişkenlerin doğrudan olarak ifade edildiği denklemlerdir.

$$ax + by = c$$

alan x_1
 ad sayısı x_2
 kumaş x_3
 fiyat y

değişken $\left. \begin{array}{l} ax_1 + bx_2 + cx_3 + b = y \\ \downarrow \text{ağırlık} \quad \text{biyar} \end{array} \right\}$

taban fiyatı m. için.

Kalkülüs

• Türev, integral

Türev: Bir fonksiyonun belirli bir noktadaki anlık değişim oranını ifade eder. Özellikle gradyan gibi optimizasyon teknikleri, bir fonksiyonun türevini kullanarak en uygun parametre değerlerini bulmaya yarar.

Integral: Bir fonksiyonun toplamını alır ve belirli bir aralıktaki sınırlar arasındaki toplamı hesaplar için kullanılır.

hız $\xleftrightarrow[\text{türev}]{\text{integral}}$ konum

olasılık ve istatistik

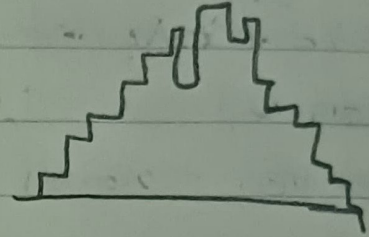
day : Bir deneyde gözlemlenebilecek sonuçlardan
herbiridir.

örnek uzay : Olası tüm sonuçların oluşturdugu
kümelerdir. $S = \{4, 7\}$

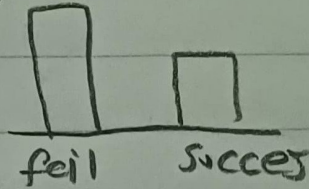
Koşullu olasılık : Bir olayın başka bir olayın
gerçekleşme durumuna bağlı olarak gerçekleşme
olasılığıdır.

olasıtsel Dağılımlar

Normal Dağılım (Gauss Dağılımı)



Bernoulli Dağılım



izlenimli olaylar.

Binom Dağılım



3 defa para atcam

ortada 3 defa tura gelme olasılığı gibi

Poisson Dağılımı : Belirli bir zaman aralığında nadir
gerçekleşen olayların sayısının modellenmesi için kullanılır.
Örneğin bir süre boyunca trafik kazalarında gerçekleşen
bölge kazası sayısını verir.

Koşullu olasılık :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Temel İstatistiksel Özellikler.

→ Ortalama (mean)

→ Medyan (median) : Ortadaki sayı

→ Mod : Tezkarane sayı

→ Varyans : Verilerin ortalamadan ne kadar ortalama aldığına
ölçüsüdür. Değişkenliği ifade eder.

Standard Sapma Varyans Korelasyon alıncak hesaplanır Ortalamadan ne kadar saptığı hakkında bilgi verir

⑤ **Gözetimli Öğrenme (Supervised Learning)**
Eğitim verilerinde girdi ve çıktı arasındaki ilişkiyi öğrenen amaçlayan bir ml yerleşimidir.
* Girdiler feature çıktıları ise etiketler temsil eder

Adımları

1) Veri toplama ve Hazırlığı

2) Model seçimi ve eğitimi

3) Model Değerlendirmesi

4) Model Ayarlaması ve Hiperparametre Optimizasyonu

* Sınıflandırma ve regresyon gibi çeşitli problemleri çözmek için kullanılır.

* Girdiler (feature) ve çıktıları (etiketler) izni de ml eğitme için kullanılabilir.

Sınıflandırma

• K-En Yakın Komşu ✓

• Karar ağaçları ✓

• Rasgele Orman ✓

✓ • Logistic Regresyon

✓ • Destek Vektör makinesi

• Naive Bayes

Regresyon

• Lineer Regresyon

• Çoğlu Lineer Regresyon

• Polinom Regresyon