

16.06.2025

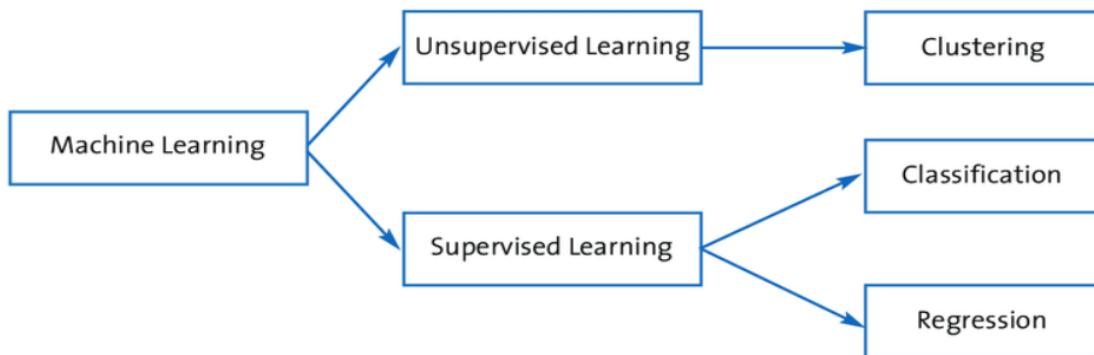
Metrikler

Model Performansını
Değerlendirmek

Esra Aydin
LOTUS AI



MACHINE LEARNING



Makine öğrenimi algoritmaları, geçiş verilerden öğrenerek benzer durumlar için sonuç tahmin etmek amacıyla kullanılır. Bu algoritmalar, problemin türüne göre yaklaşımalar içerir:

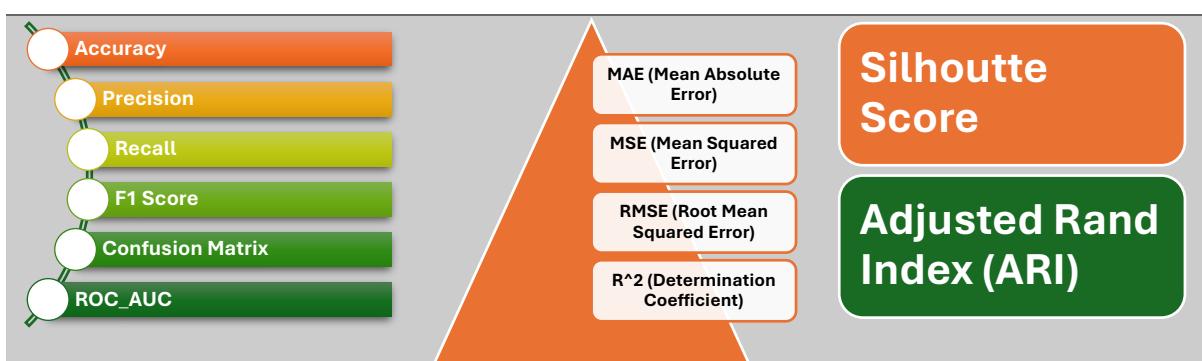
- **Regression**, değişkenler arasındaki ilişkiyi analiz ederek, sayısal bir çıktıyı tahmin etmeyi amaçlar. Örneğin, bir evin büyüklüğüne göre fiyatını thamin etmek.
- **Classification**, verilere girdiye göre verinin ait olduğu sınıfı belirler. Çıktılar ayrık ve sonlu değerlerdir. Örneğin, bir e-postanın 'spam' veya 'spam değil' olarak etiketlenmesi.
- **Clustering**, Denetimsiz öğrenme yöntemlerinden biridir ve verileri benzerliklerine göre gruptara (kümelere) ayırmayı amaçlar. Bu yöntemler veriler önceden etiketlenmemiştir; yani hangi grubun ne olduğunu model bilmez. Model, yalnızca veri içindeki örüntülerini ve benzerlilerini analiz ederek benzer özelliklere sahi verileri aynı kümeye toplar. Örneğin, bir müşteri veri setinde, alışveriş alışkanlıklarına göre müşterilerin segmentlere ayrılması (örneğin, "sadık müşteri", "tek seferlik alışveriş yapan", "fiyat odaklı müşteri" gibi) kümleme ile yapılabilir.

METRİKLER

1.1 Sınıflandırma Metrikleri:

1.2 Regresyon Metrikleri:

1.3 Clustering Metrikleri



SINIFLANDIRMA MODELLERİNİN DEĞERLENDİRİLMESİ:

Confusion Matrix (Karmaşıklık Matrisi): TP, TN, FP, FN değerlerini tablo şeklinde gösterir.

- **True Positive (TP):** Doğru şekilde tahmin edilen pozitif durumlar.
- **False Negative (FN):** Kaçırılan pozitif durumlar (yanlışlıkla negatif olarak tahmin edilmiş).
- **True Negative (TN):** Doğru şekilde tahmin edilen negatif durumlar.
- **False Positive (FP):** Yanlış bir şekilde pozitif olarak tahmin edilen negatif örneklerdir.

Örneğin:

- True Positive (TP): Bir müşterinin ürün alıp almayacağı tahmin eden bir modelde, TP, ürün alacağını tahmin ettiği ve gerçekten ürün alan müşterileri temsil eder.
- False Positive (FP): Bir müşterinin ürün alıp almayacağı tahmin eden bir modelde, FP, ürün alacağını tahmin ettiği ancak aslında ürün almayan müşterileri temsil eder.
- True Negative (TN): Bir müşterinin ürün alıp almayacağı tahmin eden bir modelde, TN, ürün almayacağı tahmin ettiği ve gerçekten de ürün almayan müşterileri temsil eder.
- False Negative (FN): Bir müşterinin ürün alıp almayacağı tahmin eden bir modelde, FN, ürün almayacağı tahmin ettiği ancak aslında ürün alan müşterileri temsil eder.

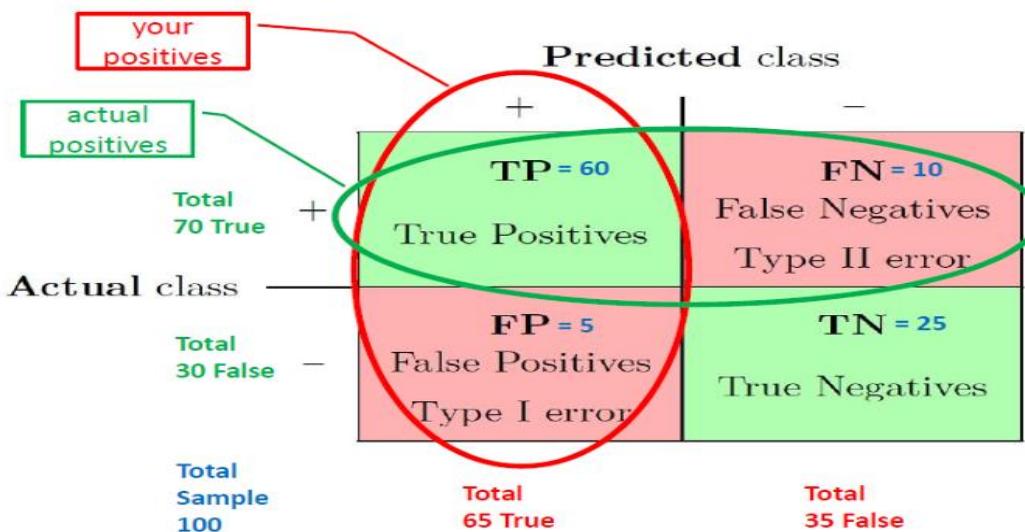


Figure 1

Accuracy: Doğru olarak sınıflandırılan örneklerin yüzdesidir. $\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$

Recall: pozitif olarak tahmin etmemiz gereken işlemlerin ne kadarını pozitif olarak tahmin ettiğimizi gösteren bir metriktir.'Gerçek pozitiflerin ne kadarı doğru bir şekilde tanımlandı.'

$$\text{Recall} = \frac{TP}{TP+FN}$$

Precision: pozitif olarak tahminlediğimiz değerlerin gerçekten kaç adedinin pozitif olduğunu göstermektedir. $\text{Precision} = \frac{TP}{TP+FP}$

F1-Score: Precision ve Recall harmonik ortalamasıdır. $F1 - Score = \frac{2 * Precision * Recall}{Recall+Precision}$

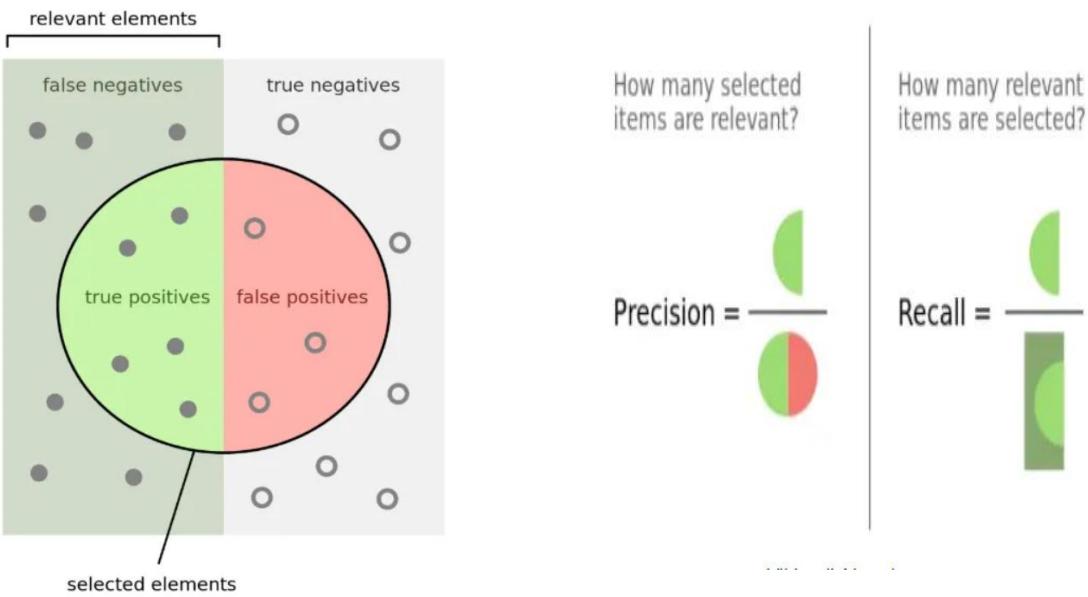


Figure 1'e göre işlemler:

$$\text{Accuracy: } 60 + 25 / 100 = 0.85$$

$$\text{Recall: } 60 / (60 + 10) \approx 0.86$$

$$\text{Precision: } 60 / (60 + 5) \approx 0.92$$

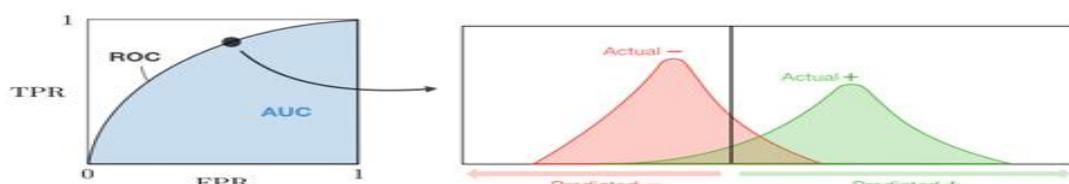
$$\text{F1-Score: } (2 * 0.86 * 0.92) / (0.86 + 0.92) \approx 0.89$$

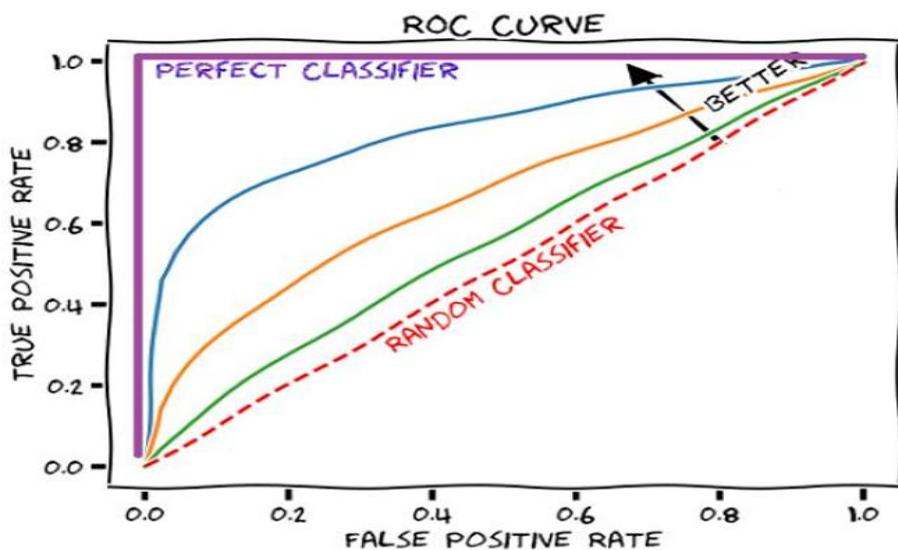
Düzen Performans Metrikleri

- **True Negative Rate (TNR):** Doğru tahmin edilen negatif durumlar, aynı zamanda Spesifiklik (Specificity) olarak da adlandırılır. Yani modelin negatif sınıfları doğru tahmin etme yeteneğini gösterir. $TNR = \frac{TN}{TN+FP}$
- **False Positive Rate (FPR):** Yanlış pozitiflerin gerçek negatiflere oranı. Yani modelin negatif olanları yanlışlıkla pozitif olarak tahmin etme oranıdır. $FPR = \frac{FP}{FP + TN}$
- $FPR = 1 - TNR$

ROC Eğrisi ve AUC:

- **ROC Eğrisi:** Farklı eşik değerleri için True Positive Rate (TPR) ile False Positive Rate (FPR)'yi karşılaştırılan bir grafiktir. Bu, duyarlılık ve özgüllük arasındaki ticaret ilişkisini görselleştirmeye yardımcı olur.
- **AUC (Area Under the Curve):** ROC eğrisinin altındaki alan, modelin sınıflar arasında ayırt etme yeteneğini ölçer.





REGRESYON MODELLERİNİN DEĞERLENDİRİLMESİ:

MAE	MSE	RMSE	R ²
$\frac{\sum_{i=1}^n y_i - \hat{y}_i }{n}$	$\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$	$\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2}$	$1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2}$

MAE (Mean Absolute Error) – Ortalama Mutlak Hata: Gerçek değerler ile tahmin edilen değerler arasındaki farkların mutlak değerlerinin ortalamasıdır.

MSE (Mean Squared Error) – Ortalama Kare Hata: Gerçek değerler ile tahmin değerleri arasındaki farkların **karelerinin** ortalamasıdır.

RMSE (Root Mean Squared Error) – Karekök Ortalama Kare Hata: MSE'nin kareköküdür.

R² (R-Squared) – Determinasyon Katsayısı: Modelin, toplam değişkenliğin ne kadarını açıkladığını gösterir.

Örnek:

GÖZLEM (i)	GERÇEK DEĞER Yİ	TAHMİN DEĞERİ Yİ
1	3	2.5
2	5	4.8
3	2	2.9
4	7	6.6
5	4	3.5

1. MAE Hesabı (Ortalama Mutlak Hata)

$$MAE = 1/5(|3-2.5| + |5-4.8| + |2-2.9| + |7-6.6| + |4-3.5|) = 1/5(0.5 + 0.2 + 0.9 + 0.4 + 0.5) = 2.5/5 = 0.5$$

- ♦ Ortalama tahmin hatası 0.5 birimdir.



2. MSE Hesabı (Ortalama Kare Hata)

$$\begin{aligned} \text{MSE} &= 1/5((3-2.5)^2 + (5-4.8)^2 + (2-2.9)^2 + (7-6.6)^2 + (4-3.5)^2) \\ &= 1/5(0.25 + 0.04 + 0.81 + 0.16 + 0.25) = 1.51/5 = 0.302 \end{aligned}$$

- ◆ Hataların karelerinin ortalaması 0.302

3. RMSE Hesabı (Karekök Ortalama Kare Hata)

$$\text{RMSE} = 0.302 \approx 0.55$$

- ◆ Ortalama hata büyüklüğü yaklaşık 0.55 birim.

4. R² Hesabı (Determinasyon Katsayısı)

Önce gerçek değerlerin ortalamasını bulalım:

$$y^- = (3+5+2+7+4) / 5 = 21/5 = 4.2$$

Toplam değişkenlik (Total Sum of Squares):

$$\begin{aligned} \sum(y_i - y^-)^2 &= (3 - 4.2)^2 + (5 - 4.2)^2 + (2 - 4.2)^2 + (7 - 4.2)^2 + (4 - 4.2)^2 = 1.44 + 0.64 + 4.84 \\ &+ 7.84 + 0.04 = 14.8 \end{aligned}$$

Model hata kareleri toplamı (Residual Sum of Squares) = 1.51 (yukarıda MSE hesabında toplamı)

$$R^2 = 1 - (1.51) / (14.8) \approx 1 - 0.102 = 0.898$$

- ◆ Model verideki değişkenliğin yaklaşık %89.8'ini açıklıyor.