

Lojistik Regresyon – Performans Metrikleri

Atil Samancioglu

1 Giriş

Lojistik regresyon, sınıflandırma problemlerinde kullanılan güçlü bir algoritmadır. Bu algoritmanın başarımını ölçmek için doğrusal regresyondan farklı metriklere ihtiyaç duyulur. Bu dokümanda özellikle **binary classification** problemleri için kullanılan *Confusion Matrix*, *Accuracy*, *Precision*, *Recall* ve *F1/F_βcore* gibi metrikler açıklanacaktır.

2 Confusion Matrix (Karmaşıklık Matrisi)

Confusion matrix, sınıflandırma problemlerinde modelin doğru ve yanlış tahminlerini dört kategoriye ayırır:

- **True Positive (TP)**: Gerçek değer pozitif, tahmin de pozitif.
- **True Negative (TN)**: Gerçek değer negatif, tahmin de negatif.
- **False Positive (FP)**: Gerçek değer negatif, tahmin pozitif.
- **False Negative (FN)**: Gerçek değer pozitif, tahmin negatif.

Confusion Matrix

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

Figure 1: Confusion Matrix bileşenleri.

Accuracy (Doğruluk)

doğru tahmin yaptıklarımızın tümüne oranı

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$


Genel olarak modelin ne kadar doğru tahmin yaptığını gösterir. Ancak **dengesiz veri kümelerinde** yanıltıcı olabilir.

3 Precision ve Recall

Precision (Kesinlik)

pozitif tahminlerimi oranlar ve benim için pozitif tahminlerin önemi fazla ise bu metrice bakılır

$$\text{Precision} = \frac{TP}{TP + FP}$$

 azaltmaya çalışırız

Modelin pozitif tahminleri arasında, kaç tanesinin gerçekten pozitif olduğunu ölçer.

Recall (Duyarlılık)

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

 azaltmaya çalışırız

Gerçek pozitif örneklerden kaç tanesini modelin doğru tahmin ettiğini gösterir.

Kullanım Senaryoları

Precision'ın önemli olduğu durumlar:

- Spam sınıflandırma: Gerçek spam olmayan e-postaların yanlışlıkla spam olarak işaretlenmemesi gerekir. FP azaltılmalı → Precision yüksek olmalı.

Recall'ın önemli olduğu durumlar:

- Hastalık teşhisi (örneğin diyabet): Hasta bir bireyin hasta değil olarak sınıflandırılması ciddi bir hatadır. FN azaltılmalı → Recall yüksek olmalı.

4 F_1 ve F_β skorları

Precision ve Recall metrikleri arasında denge sağlamak için kullanılır.

F1 Skoru

$$F_1 = \frac{2 \cdot (\text{Precision} \cdot \text{Recall})}{\text{Precision} + \text{Recall}}$$

F_β skoru

$$F_\beta = (1 + \beta^2) \cdot \frac{\text{Precision} \cdot \text{Recall}}{(\beta^2 \cdot \text{Precision}) + \text{Recall}}$$

- $\beta = 1$: Precision ve Recall eşit ağırlıkta (F1).
- $\beta < 1$: Precision daha önemli (örnek: $F_{0.5}$).
- $\beta > 1$: Recall daha önemli (örnek: F_2).

F-skoru daima 0 ile 1 arasında bir değer alır. Bu değer 1'e yaklaştıkça modelin hem **precision** hem de **recall** açısından güçlü olduğu anlaşılır. F -skorunun 0'a yaklaşması ise modelin bu iki metriktan en az birinde ciddi sorunlar yaşadığını gösterir.

5 Sonuç

Logistic regresyon modellerinin başarımını anlamak için sadece doğruluk (accuracy) yeterli olmayabilir. Özellikle dengesiz veri kümelerinde precision, recall ve $F1/F_\beta$ skorları oldukça önemlidir. Hangi metriğin kullanılacağı, probleme ve risklere göre değişir.