



Model Performansının Değerlendirilmesi

YZL-455 Makine Öğrenmesi



- Karışıklık (confusion) matrisi veri kümesindeki gerçek sınıf etiketi ile öngörülen sınıf etiketi sayılarını içerir. İki sınıflı bir veri kümesi için örnek karışıklık matrisi aşağıdaki gibidir.

		Gerçek	
		YES	NO
Tahmin	YES	True Positive <ul style="list-style-type: none">Kurt vardır.Çoban "Kurt geldi!" der.Çoban kahraman ilan edilir. TP sayısı: 1	False Positive <ul style="list-style-type: none">Kurt yoktur.Çoban "Kurt geldi!" der.Köylüler çobana kızar. FP sayısı: 1
	NO	False Negative <ul style="list-style-type: none">Kurt vardır.Çoban "Kurt yok" der.Kurt kuzuları yer. FN sayısı: 8	True Negative <ul style="list-style-type: none">Kurt yoktur.Çoban "Kurt yok" der.Herkes iyi. TN sayısı: 90

- **DP**(doğru pozitif) / **TP**(true positive):

Gerçek durumu **pozitifken** öngörü sonucu da **pozitif** çıkan örneklerin sayısı. (Model Doğru öngörü yapmıştır.)

- **YN**(yanlış negatif) / **FN**(false negative):

Gerçek durumu **pozitifken** öngörü sonucu **negatif** çıkan örneklerin sayısı. (Model Yanlış öngörü yapmıştır.)

- **DN**(doğru negatif) / **TN**(true negative):

Gerçek durumu **negatifken** öngörü sonucu da **negatif** çıkan örneklerin sayısı. (Model Doğru öngörü yapmıştır.)

- **YP**(yanlış pozitif) / **FP**(false positive):

Gerçek durumu **negatifken** öngörü sonucu **pozitif** çıkan örneklerin sayısı. (Model Yanlış öngörü yapmıştır.)

Bir makine öğrenme modelinin performansını belirlemek üzere bazı ölçütlerden yararlanılır. Bu ölçütlerden en çok tercih edilenleri şunlardır:

- Doğruluk (Accuracy)
- Duyarlılık (Sensitivity-Recall)
- Özgünlük-Belirleyicilik (Specificity)
- Kesinlik (Precision)
- F score

Bu ölçütler test veri kümesi kullanılarak hesaplanır.

- Doğruluk, modelin doğru öngördüğü toplam örnek sayısının (gerçek durumu pozitifken öngörü sonucu da pozitif çıkan örneklerin sayısı + gerçek durumu negatifken öngörü sonucu da negatif çıkan örneklerin sayısı), veri kümesindeki toplam örnek sayısına oranıdır.
- Pozitif ve negatif örneklerin ne kadarının model tarafından doğru öngörülebildiğini gösterir.

$$\text{Doğruluk (Accuracy)} = (DP+DN) / (DP+DN+YP+YN)$$

$$\text{Doğruluk (Accuracy)} = (DP+DN) / (m)$$

- Özellikle dengesiz veri kümelerinde sadece Doğruluk (Accuracy) metriğine göre yorum yapmak yanıltıcı olabilir. Bu durumlarda ASLA bir ölçüm olarak kullanılmamalıdır.
- Örneğin 2 sınıflı bir veri kümesi olduğunu varsayalım.
- Sınıf 1 test örneklerinin sayısı= 9990
- Sınıf 2 test örneklerinin sayısı= 10
- Eğer model bütün test örneklerini Sınıf 1 olarak öngörürse
- Doğruluk= $9990/10000 = \%99.9$ olur.
- **Sınıf 2 örneklerinden hiç biri öngörülmediği halde doğruluk bilgisi bu durumda yanıltıcı olabilir !!!.**

Hata oranı (Error rate)

- Makine öğrenme modelinin hata oranı (error rate) aşağıdaki gibi hesaplanmaktadır.

$$\text{Hata oranı (Error rate)} = (YP + YN) / (DP + DN + YP + YN)$$

$$\text{Hata oranı (Error rate)} = 1 - \text{Doğruluk (Accuracy)}$$

- Doğru Pozitif Oranı;
- Modelin, gerçek pozitif örnekler içinden doğru öngörülen pozitif örnekleri öngörmedeki etkinliği duyarlılık (sensitivity, true positive rate) olarak adlandırılmaktadır.
- Kısaca, modelin veri setindeki pozitif sınıfı öngörmedeki etkinliğidir.
- Pozitif örneklerden hangi oranda model tarafından doğru öngörülebildiğini gösterir.

- Duyarlılık, modelin doğru öngördüğü pozitif örnek sayısının, toplam gerçek pozitif örnek sayısına (gerçek durumu pozitifken öngörü sonucu da pozitif çıkan örneklerin sayısı + gerçek durumu pozitifken öngörü sonucu negatif çıkan örneklerin sayısı) oranıdır.

Duyarlılık (Sensitivity) = DP/Gerçek Pozitif sayısı

Duyarlılık (Sensitivity) = $DP / (DP + YN)$

- Doğru Negatif Oranı;
- Modelin, gerçek negatif örnekler içinden doğru öngörülen negatif örnekleri öngörmedeki etkinliği özgünlük (specificity, true negative rate) biçiminde adlandırılmaktadır.
- Kısaca, modelin veri setindeki negatif sınıfı öngörmedeki etkinliğidir.
- Negatif örneklerin hangi oranında model tarafından doğru öngörülebildiğini gösterir.

- Özgünlük, doğru öngörülen negatif örnek sayısının, toplam gerçek negatif örnek sayısına (gerçek durumu negatifken öngörü sonucu da negatif çıkan örneklerin sayısı + gerçek durumu negatifken öngörü sonucu pozitif çıkan örneklerin sayısı) oranıdır.

Özgünlük-Belirleyicilik (Specificity) = $DN / \text{Gerçek Negatif sayısı}$

Özgünlük-Belirleyicilik (Specificity) = $DN / (DN + YP)$

- Pozitif Öngörü Değer (Positive Predictive Value);
- Toplam öngörülen pozitif örneklerin dışındaki pozitif örneklerin oranı.
- Modelin, doğru olarak öngördüğü pozitif örnek sayısının, öngörülen toplam pozitif örnek sayısına (gerçek durumu pozitifken öngörü sonucu da pozitif çıkan örneklerin sayısı + gerçek durumu negatifken öngörü sonucu pozitif çıkan örneklerin sayısı) oranıdır.
- Kesinlik (Precision) = $DP / \text{Öngörülen pozitif sayısı}$
- Kesinlik (Precision) = $DP / (DP + YP)$
- Duyarlılık (Sensitivity) ve Kesinlik (Precision) arasında ters orantı vardır

- Model performansı için hem kesinlik (precision) hem de duyarlılık (sensitivity-recall) ölçütlerine göre optimizasyon yapmak gerekir.
- F1-Score, iki ölçüt yerine modelin tek ölçüt ile değerlendirilmesi için kullanılır. F1-Score değeri kesinlik ve duyarlılık arasında bir yerde salınır.
- F1-Score, kesinlik (precision) ve duyarlılık (sensitivity-recall) değerlerinin harmonik ortalaması (harmonic mean) alınarak hesaplanır.

Sensitivity = Recall= true positive rate = $DP / (DP+YN)$

Precision = $DP / (DP +YP)$

F1–Score = $2*(Precision*Recall) / (Precision + Recall)$

F1-Score ne kadar yüksek olursa modelin performansı o kadar iyidir.

- Doğruluk (Accuracy)= $(DP+DN)/(DP+YP+YN+DN)$
 $= (75+420)/(75+300+5+420)=0,62$
(Model %62 oranında doğru öngörü yapabilmekte)
- Duyarlılık (Sensitivity)= $DP/(DP+YN) = 75/80 = 0,94$
(Sadece pozitif sınıf değerlerini içeren örnekler göz önüne alındığında modelin öngörü başarısı %94. Bu değer pozitif sınıfının öngörülmesindeki başarıyı gösterir.)
- Özgünlük-Belirleyicilik (Specificity)= $DN/(YP+DN) = 420/720 = 0,58$
(Sadece negatif sınıf değerlerini içeren örnekler göz önüne alındığında modelin öngörü başarısı %58. Bu değer negatif sınıfının öngörülmesindeki başarıyı gösterir.)
- Kesinlik (Precision) = $DP / (DP+YP) = 75 / 75+300 = 0,2$
(Model pozitif sınıfı için %20 oranında isabetli öngörü yapabilmekte)

		Öngörülen		
		Pozitif	Negatif	Toplam
Gerçek	Pozitif	75 (DP)	5 (YN)	80
	Negatif	300 (YP)	420 (DN)	720
	Toplam	375	425	800

Accuracy=?

Precision=?

Recall-Sensitivity=?

F1 score=?

TP 130	FP 20
FN 25	TN 125

3x3 karmaşıklık matrisi

		Predicted		
		Greyhound	Mastiff	Samoyed
Actual	Greyhound	P_{GG}	P_{MG}	P_{SG}
	Mastiff	P_{GM}	P_{MM}	P_{SM}
	Samoyed	P_{GS}	P_{MS}	P_{SS}

3x3 karmaşıklık matrisi

- **PGG** : Gerçek durumu Greyhound, öngörü sonucu da Greyhound çıkan örneklerin sayısı.
- **PMG** : Gerçek durumu Greyhound, öngörü sonucu Mastiff çıkan örneklerin sayısı.
- **PSG** : Gerçek durumu Greyhound, öngörü sonucu Samoyed çıkan örneklerin sayısı.
- **PGM** : Gerçek durumu Mastiff, öngörü sonucu Greyhound çıkan örneklerin sayısı.
- **PMM** : Gerçek durumu Mastiff, öngörü sonucu Mastiff çıkan örneklerin sayısı.
- **PSM** : Gerçek durumu Mastiff, öngörü sonucu Samoyed çıkan örneklerin sayısı.

- PGS : Gerçek durumu Samoyed, öngörü sonucu Greyhound çıkan örneklerin sayısı.
- PMS : Gerçek durumu Samoyed, öngörü sonucu Mastiff çıkan örneklerin sayısı.
- PSS : Gerçek durumu Samoyed, öngörü sonucu Samoyed çıkan örneklerin sayısı.

Doğru Pozitifler (True Positives)

- Greyhound sınıfı için Doğru Pozitifler (True Positives), karışıklık matrisindeki PGG değişkenidir.
- Mastiff sınıfı için Doğru Pozitifler (True Positives) karışıklık matrisindeki PMM değişkenidir.
- Samoyed sınıfı için Doğru Pozitifler, (True Positives) karışıklık matrisindeki PSS değişkenidir.

3x3 karmaşıklık matrisi

- Doğru Negatifler (True Negatives)
- Greyhound sınıfı için Doğru Negatifler :
- $DN (TN) = PMM + PSM + PMS + PSS$

		Predicted		
		Greyhound	Mastiff	Samoyed
Actual	Greyhound			
	Mastiff		P_{MM}	P_{SM}
	Samoyed		P_{MS}	P_{SS}

3x3 karmaşıklık matrisi

- Doğru Negatifler (True Negatives)
- Mastiff sınıfı için Doğru Negatifler :
- $DN (TN) = P_{GG} + P_{SG} + P_{GS} + P_{SS}$

		Predicted		
		Greyhound	Mastiff	Samoyed
Actual	Greyhound	P_{GG}		P_{SG}
	Mastiff			
	Samoyed	P_{GS}		P_{SS}

3x3 karmaşıklık matrisi

- Doğru Negatifler (True Negatives)
- Samoyed sınıfı için Doğru Negatifler :
- $DN (TN) = P_{GG} + P_{MG} + P_{GM} + P_{MM}$

		Predicted		
		Greyhound	Mastiff	Samoyed
Actual	Greyhound	P_{GG}	P_{MG}	
	Mastiff	P_{GM}	P_{MM}	
	Samoyed			

3x3 karmaşıklık matrisi

- Yanlış Pozitifler (False Positives)
- Greyhound sınıfı için Yanlış Pozitifler :
- $YP (FP) = P_{GM} + P_{GS}$

		Predicted		
		Greyhound	Mastiff	Samoyed
Actual	Greyhound	P_{GG}	P_{MG}	P_{SG}
	Mastiff	P_{GM}	P_{MM}	P_{SM}
	Samoyed	P_{GS}	P_{MS}	P_{SS}

3x3 karmaşıklık matrisi

- Yanlış Pozitifler (False Positives)
- Mastiff sınıfı için Yanlış Pozitifler:
- $YP (FP) = PMG + PMS$

		Predicted		
		Greyhound	Mastiff	Samoyed
Actual	Greyhound	P_{GG}	P_{MG}	P_{SG}
	Mastiff	P_{GM}	P_{MM}	P_{SM}
	Samoyed	P_{GS}	P_{MS}	P_{SS}

3x3 karmaşıklık matrisi

- Yanlış Pozitifler (False Positives)
- Samoyed sınıfı için Yanlış Pozitifler:
- $YP (FP) = PSG + PSM$

		Predicted		
		Greyhound	Mastiff	Samoyed
Actual	Greyhound	P_{GG}	P_{MG}	P_{SG}
	Mastiff	P_{GM}	P_{MM}	P_{SM}
	Samoyed	P_{GS}	P_{MS}	P_{SS}

3x3 karmaşıklık matrisi

- Yanlış Negatifler (False Negatives)
- Greyhound sınıfı için Yanlış Negatifler:
- $YN (FN) = PMG + PSG$

		Predicted		
		Greyhound	Mastiff	Samoyed
Actual	Greyhound	P_{GG}	P_{MG}	P_{SG}
	Mastiff	P_{GM}	P_{MM}	P_{SM}
	Samoyed	P_{GS}	P_{MS}	P_{SS}

3x3 karmaşıklık matrisi

- Yanlış Negatifler (False Negatives)
- Mastiff sınıfı için Yanlış Negatifler:
- $YN (FN) = P_{GM} + P_{SM}$

		Predicted		
		Greyhound	Mastiff	Samoyed
Actual	Greyhound	P_{GG}	P_{MG}	P_{SG}
	Mastiff	P_{GM}	P_{MM}	P_{SM}
	Samoyed	P_{GS}	P_{MS}	P_{SS}

3x3 karmaşıklık matrisi

- Yanlış Negatifler (False Negatives)
- Samoyed sınıfı için Yanlış Negatifler:
- $YN (FN) = PGS + PMS$

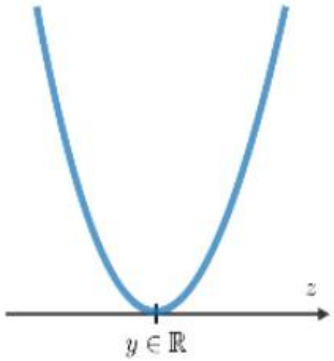
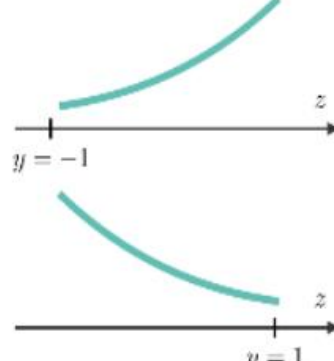
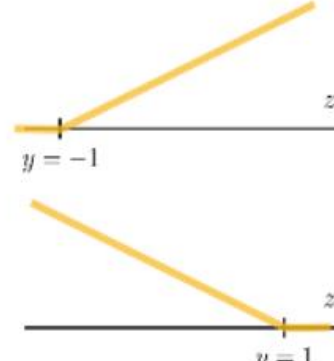
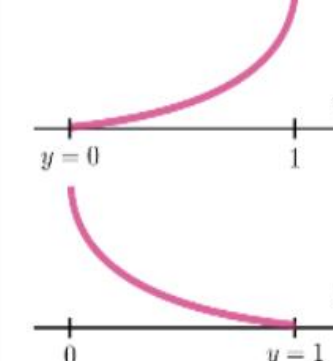
		Predicted		
		Greyhound	Mastiff	Samoyed
Actual	Greyhound	P_{GG}	P_{MG}	P_{SG}
	Mastiff	P_{GM}	P_{MM}	P_{SM}
	Samoyed	P_{GS}	P_{MS}	P_{SS}

- Doğruluk (Accuracy) = $(PGG + PMM + PSS) / (PGG + PMG + PSG + PGM + PMM + PSM + PGS + PMS + PSS)$
- Kesinlik (Precision)
- Greyhound sınıfı için Precision:
- Precision (G) = $TP / (TP + FP) = PGG / (PGG + (PGM + PGS))$
- Mastiff sınıfı için Precision :
- Precision (M) = $TP / (TP + FP) = PMM / (PMM + (PMG + PMS))$
- Samoyed sınıfı için Precision :
- Precision (S) = $TP / (TP + FP) = PSS / (PSS + (PSG + PSM))$

- Duyarlılık (Sensitivity - True Positive Rate)
- Greyhound sınıfı için Sensitivity :
- Sensitivity (G) = $TP / (TP + FN) = PGG / (PGG + (PMG + PSG))$
- Mastiff sınıfı için Sensitivity:
- Sensitivity (M) = $TP / (TP + FN) = PMM / (PMM + (PGM + PSM))$
- Samoyed sınıfı için Sensitivity:
- Sensitivity (S) = $TP / (TP + FN) = PSS / (PSS + (PGS + PMS))$

- Özgünlük (Specificity - True Negative Rate)
- Greyhound sınıfı için Specificity :
 - Specificity (G) = $TN / (TN + FP) = (PMM + PSM + PMS + PSS) / ((PMM + PSM + PMS + PSS) + (PGM + PGS))$
- Mastiff sınıfı için Specificity :
 - Specificity (M) = $TN / (TN + FP) = (PGG + PSG + PGS + PSS) / ((PGG + PSG + PGS + PSS) + (PMG + PMS))$
- Samoyed sınıfı için Specificity :
 - Specificity (S) = $TN / (TN + FP) = (PGG + PMG + PGM + PMM) / ((PGG + PMG + PGM + PMM) + (PSG + PSM))$

- İyi bir öngörü yapabilen modelde kayıp değeri az olacaktır.
- Gerçek etiket değer ile model öngörüsü birebir aynı olduğu durumda kayıp 0 olacaktır.
- İyi bir makine öğrenme modelinden beklenen kayıp değerinin 0'a yakın olmasıdır.

Least squared error	Logistic loss	Hinge loss	Cross-entropy
$\frac{1}{2}(y - z)^2$	$\log(1 + \exp(-yz))$	$\max(0, 1 - yz)$	$-[y \log(z) + (1 - y) \log(1 - z)]$
			
Linear regression	Logistic regression	SVM	Neural Network

*Dünya
Seni Bekliyor!*



Dinlediğiniz için teşekkürler!

Öğr. Gör. Mustafa Sami CÜCEN

Yazılım Mühendisliği



ostimteknik.edu.tr



Ostim, 100. Yıl Blv 55/F, Yenimahalle/Ankara



(0312) 386 10 92



/ostimteknikuniv

