



Sınıflandırma Modelini Değerlendirme

- Model başarımını değerlendirme ölçütleri nelerdir?
 - Hata oranı
 - Anma
 - Kesinlik
 - F-ölçütü
- Farklı modellerin başarımı nasıl karşılaştırılır?
 - ROC



Sınıflandırma Hatası

- Sınıflandırma yöntemlerinin hatalarını ölçme
 - başarı: örnek doğru sınıfa atandı
 - hata: örnek yanlış sınıfa atandı
 - hata oranı: hata sayısının toplam örnek sayısına bölünmesi
- Hata oranı sına ma kümesi kullanılarak hesaplanır.



Model Başarımını Değerlendirme

- Model başarımını değerlendirme ölçütleri
 - modelin ne kadar doğru sınıflandırma yaptığını ölçer
 - hız, ölçeklenebilirlik gibi özellikleri değerlendirmez.
- Sınıflandırma işlemlerinde gerçek veri ile öngörülen değerleri karşılaştırmak amacıyla karışıklık matrisinden (confusion matrix) yararlanır.



Karışıklık Matrisi

- Matrisin satırlarını gerçek değerlere ilişkin sınıf etiketleri, sütunlarını ise öngörülen sınıf etiketleri oluşturur.
- Karışıklık matrisi:

	ÖNGÖRÜLEN SINIF		
DOĞRU SINIF		Sınıf=Yes	Sınıf=No
	Sınıf=Yes	a	b
	Sınıf=No	c	d

a: TP (true positive)

b: FN (false negative)

c: FP (false positive)

d: TN (true negative)

Model Başarımını Değerlendirme : Doğruluk

DOĞRU SINIF	ÖNGÖRÜLEN SINIF		
		Sınıf=Yes	Sınıf=No
	Sınıf=Yes	a (TP)	b (FN)
	Sınıf=No	c (FP)	d (TN)

■ Modelin başarımı:

$$\text{Doğruluk} = \frac{a + d}{a + b + c + d} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\text{Hata Oranı} = \frac{b + c}{a + b + c + d} = \frac{\text{FN} + \text{FP}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\text{Hata Oranı} = 1 - \text{Doğruluk}$$



Örnek

Sınıflandırıcı A

TP=25	FN=25
FP=25	TN=25

Doğruluk=%50

Sınıflandırıcı B

TP=50	FN=0
FP=25	TN=25

Doğruluk=%75

Sınıflandırıcı C

TP=25	FN=25
FP=0	TN=50

Doğruluk=%75

- Hangi sınıflandırıcı daha iyi?
 - B ve C, A'dan daha iyi bir sınıflandırıcı
 - B, C'den daha iyi bir sınıflandırıcı mı?

Model Başarımını Değerlendirme : Kesinlik

DOĞRU SINIF	ÖNGÖRÜLEN SINIF		
		Sınıf=Yes	Sınıf=No
	Sınıf=Yes	a (TP)	b (FN)
	Sınıf=No	c (FP)	d (TN)

$$\begin{aligned}\text{Kesinlik} &= \frac{\text{Doğru sınıflandırılmış pozitif örnek sayısı}}{\text{Pozitif sınıflandırılmış örneklerin sayısı}} \\ &= \frac{TP}{TP + FP}\end{aligned}$$



Model Başarımını Değerlendirme : Anma

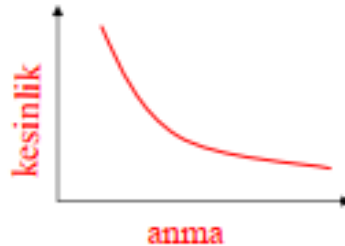
DOĞRU SINIF	ÖNGÖRÜLEN SINIF		
		Sınıf=Yes	Sınıf=No
	Sınıf=Yes	a (TP)	b (FN)
	Sınıf=No	c (FP)	d (TN)

$$\text{Anma} = \frac{\text{Doğru sınıflandırılmış pozitif örnek sayısı}}{\text{Pozitif örneklerin sayısı}}$$

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

Anma / Kesinlik

- A modeli B modelinden daha iyi anma ve kesinlik değerine sahipse A modeli daha iyi bir sınıflandırıcıdır.
- Kesinlik ve anma arasında ters orantı var.





Sınıflandırıcıları Karşılaştırma

- Doğruluk en basit ölçüt
- Kesinlik ve anma daha iyi ölçme sağlıyor
 - Model A'nın kesinliği model B'den daha iyi ancak model B'nin anma değeri model A'dan daha iyi olabilir.



Model Başarımını Değerlendirme : F-ölçütü

- F-ölçütü: Anma ve kesinliğin harmonik ortalamasını alır.

$$\text{F-ölçütü} = \frac{2 * \text{Kesinlik} * \text{Anma}}{\text{Kesinlik} + \text{Anma}}$$

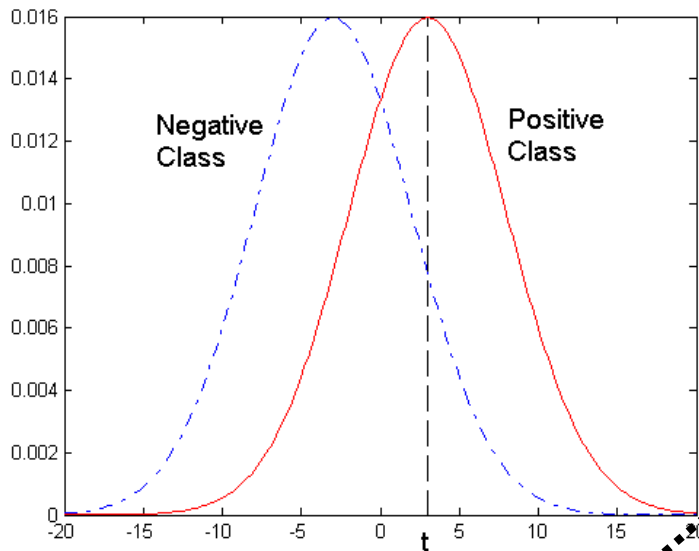


ROC (Receiver Operating Characteristic)

- İşaret işlemede bir sezicinin, gürültülü bir kanalda doğru algılama oranının yanlış alarm oranına karşı çizdirilen grafiği (algılayıcı işletim eğrisi)
- Farklı sınıflandırıcıları karşılaştırmak için ROC eğrileri
- Doğru pozitif (TPR - y eksen) oranının yanlış pozitif (FPR - x eksen) oranına karşı çizdirilen grafiği
 - $TPR = TP / (TP + FN)$
 - $FPR = FP / (TN + FP)$
- ROC üzerindeki her nokta bir sınıflandırıcının oluşturduğu bir modele karşı düşer.

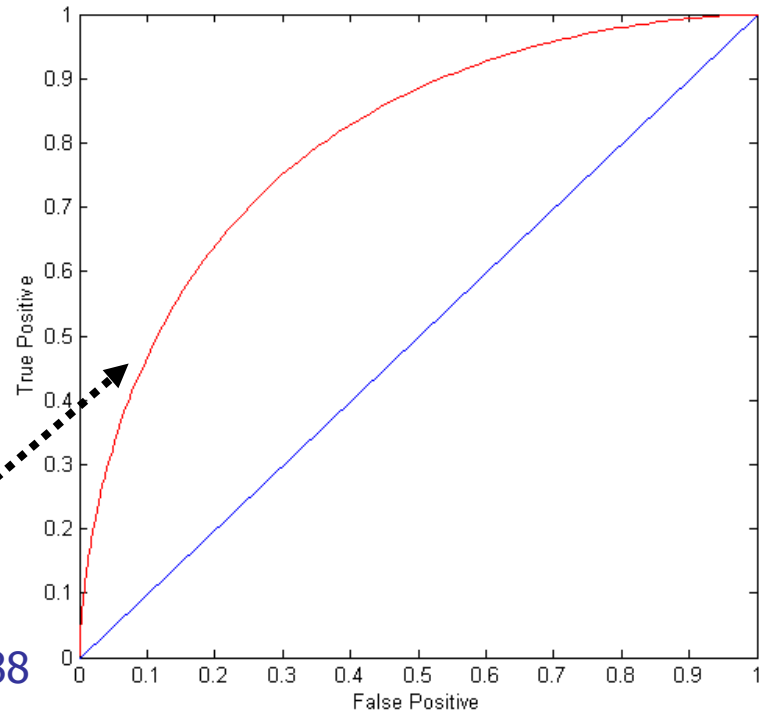
ROC Eğrisi

- iki sınıftan oluşan tek boyutlu bir veri kümesi (positive – negative)
- $x > t$ için her örnek pozitif olarak sınıflandırılıyor.



t eşik değeri için:

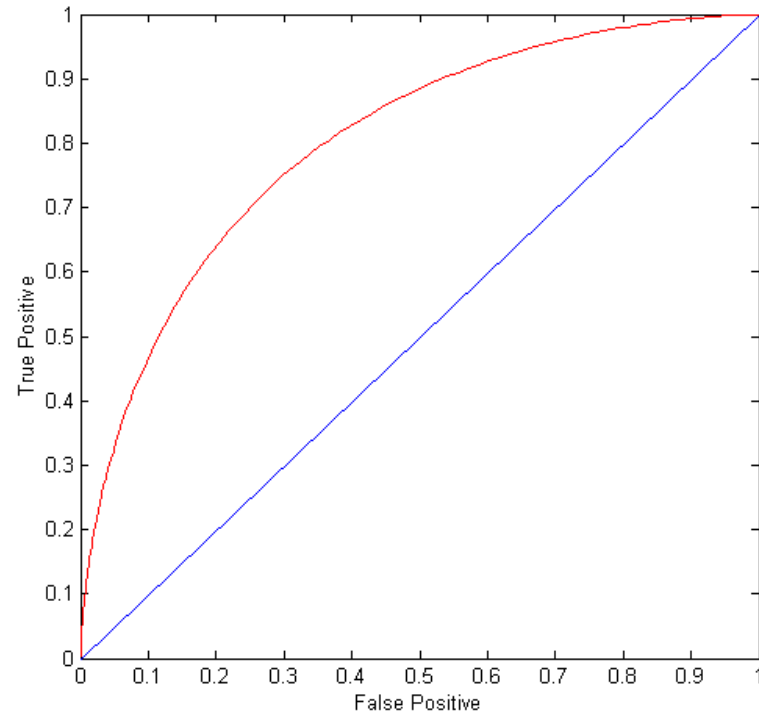
TPR=0.5, FNR=0.5, FPR=0.12, FNR=0.88



ROC Eğrisi

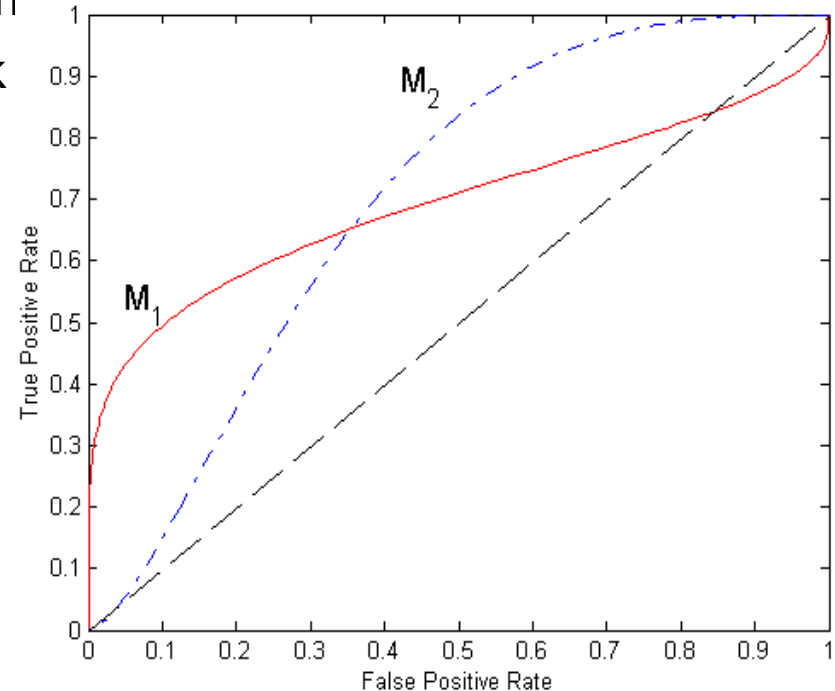
(FPR,TPR)

- (0,0): Bütün örneklerin negatif sınıflandırılması
- (1,1): Bütün örneklerin pozitif sınıflandırılması
- (0,1): ideal durum
- Çapraz çizgi:
 - Rastlantısal tahmin



ROC Eğrilerinin Kullanılması

- Farklı modelleri karşılaştırmak için
- M_1 veya M_2 birbirlerine üstünlük sağlamıyor
 - küçük FPR değerleri için M_1 daha iyi
 - büyük FPR değerleri için M_2 daha iyi
- ROC eğrisi altında kalan alan
 - ideal = 1
 - Rastlantısal tahmin=0.5



ROC Eğrisinin Çizilmesi

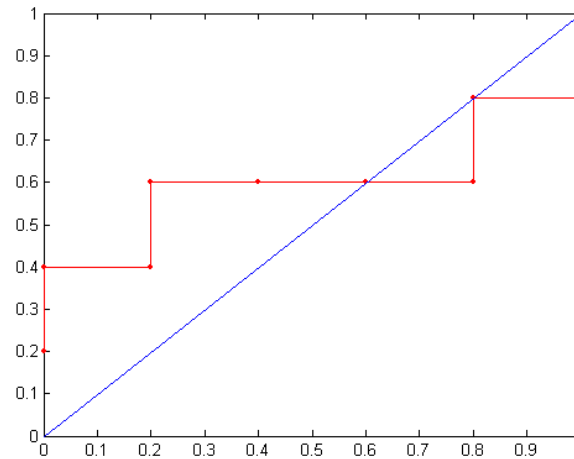
- Her örnek için $P(+|A)$ olasılığı hesaplanır.
- $P(+|A)$ değeri azalarak sıralanır.
- Her farklı $P(+|A)$ değeri için bir eşik değeri uygulanır.
- Her eşik değeri için TP, FP, TN, FN hesaplanır.

Örnek	$P(+ A)$	Doğru Sınıf
1	0.95	+
2	0.93	+
3	0.87	-
4	0.85	-
5	0.85	-
6	0.85	+
7	0.76	-
8	0.53	+
9	0.43	-
10	0.25	+

ROC Eğrisinin Çizilmesi

Sınıf	+	-	+	-	-	-	+	-	+	+	
	0.25	0.43	0.53	0.76	0.85	0.85	0.85	0.87	0.93	0.95	1.00
TP	5	4	4	3	3	3	3	2	2	1	0
FP	5	5	4	4	3	2	1	1	0	0	0
TN	0	0	1	1	2	3	4	4	5	5	5
FN	0	1	1	2	2	2	2	3	3	4	5
TPR	1	0.8	0.8	0.6	0.6	0.6	0.6	0.4	0.4	0.2	0
FPR	1	1	0.8	0.8	0.6	0.4	0.2	0.2	0	0	0

ROC Eğrisi:



$$TPR = TP / (TP + FN)$$

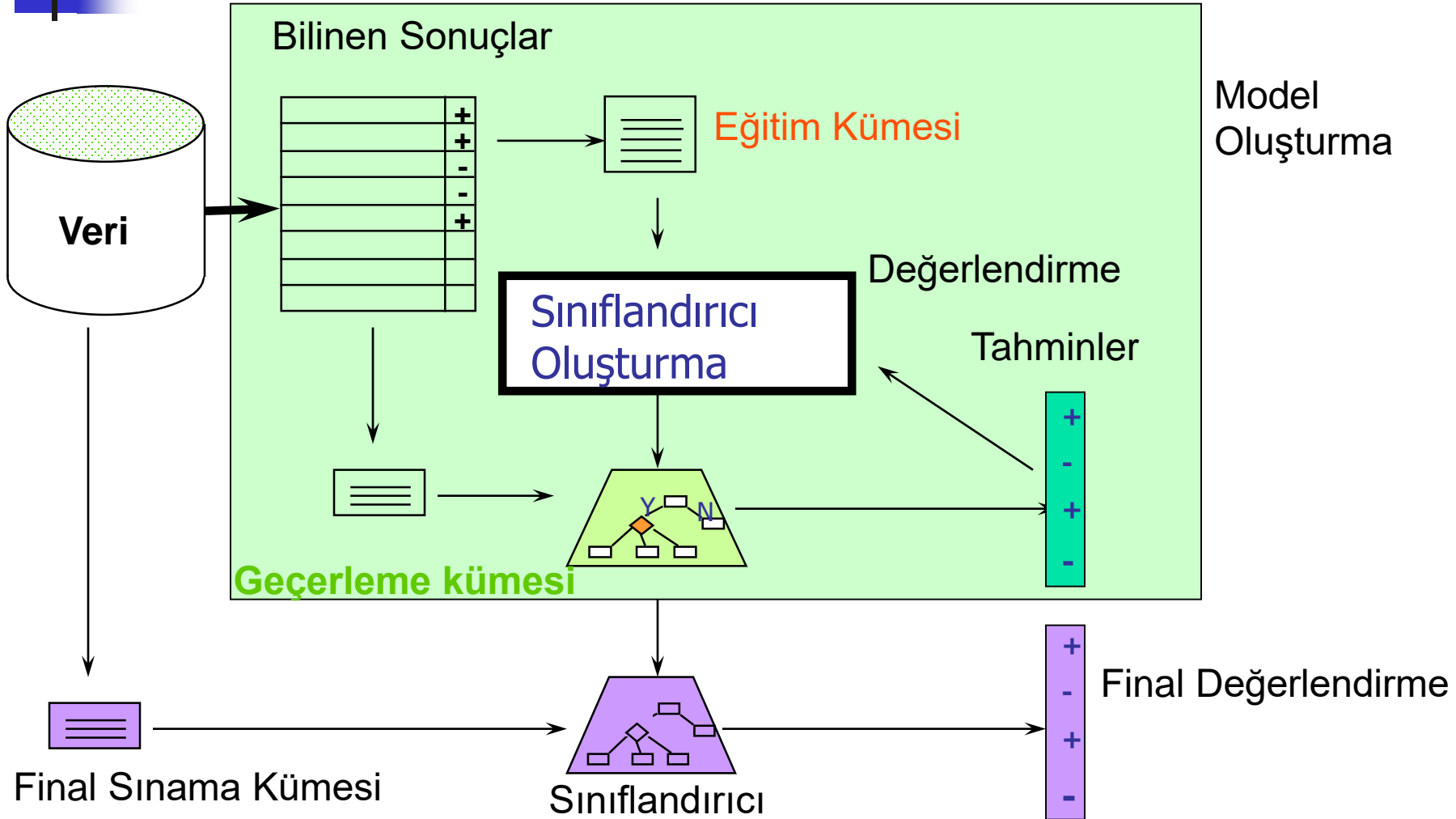
$$FPR = FP / (TN + FP)$$



Model Parametrelerini Belirleme

- Sınama kümesi sınıflandırıcı oluşturmak için kullanılmaz.
- Bazı sınıflandırıcılar modeli iki aşamada oluşturur:
 - modeli oluştur
 - parametreleri ayarla
- Sınama kümesi parametreleri ayarlamak için kullanılmaz.
- Uygun yöntem üç veri kümesi kullanma: öğrenme, geçerleme, sınama
 - geçerleme kümesi parametre ayarlamaları için kullanılır.
 - model oluşturulduktan sonra öğrenme ve geçerleme kümesi son modeli oluşturmak için kullanılabilir.

Sınıflandırma: Öğrenme, Geçerleme, Sınama



Test verisi parametre ayarlaması için kullanılmaz!



Öğrenme, sinama, geçerleme kümelerini oluşturma

- Öğrenme, sinama ve geçerleme kümelerini oluşturma
 - holdout
 - k-kat çapraz geçerleme
 - biri hariç çapraz geçerleme
 - bootstrap



Verinin Dengesiz Dağılımı

- Küçük veya dengesiz veri kümeleri için örnekler tanımlayıcı olmayabilir.
- Veri içinde bazı sınıflardan çok az örnek olabilir.
 - tıbbi veriler: %90 sağlıklı, %10 hastalık
 - elektronik ticaret: %99 alışveriş yapmamış, %1 alışveriş yapmış
 - güvenlik: %99 sahtekarlık yapmamış, %1 sahtekarlık yapmış
- Örnek: Sınıf1: 9990 örnek, Sınıf2: 10 örnek
 - bütün örnekleri sınıf1'e atayan bir sınıflandırıcının hata oranı: $9990 / 10000 = \%99,9$
 - hata oranı yanıltıcı bir ölçüt olabilir.



Dengeli Dağılım Nasıl Sağlanır?

- Veri kümesinde iki sınıf varsa
 - İki sınıfın eşit dağıldığı bir veri kümesi oluştur
 - Az örneği olan sınıftan istenen sayıda rasgele örnekler seç
 - Çok örneği olan sınıftan aynı sayıda örnekleri ekle
- Veri kümesinde iki sınıftan fazla sınıf varsa
 - Öğrenme ve sinama kümesini farklı sınıflardan aynı sayıda örnek olacak şekilde oluştur.

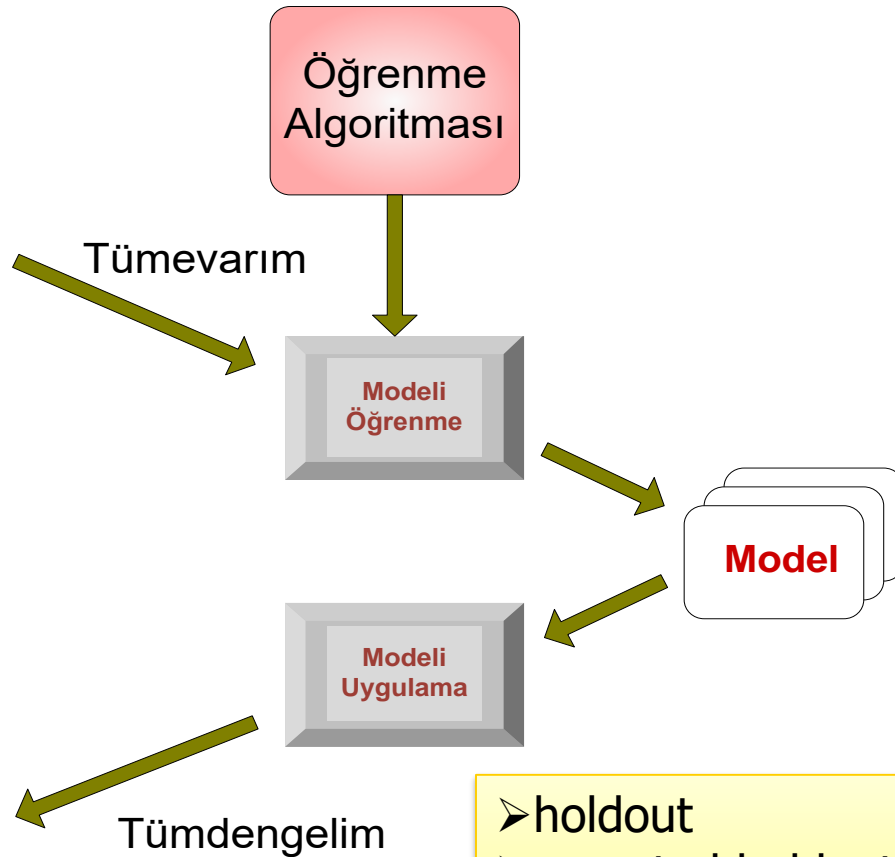
Örnek

Tid	Attrib1	Attrib2	Attrib3	Class
1	Yes	Large	125K	No
2	No	Medium	100K	No
3	No	Small	70K	No
4	Yes	Medium	120K	No
5	No	Large	95K	Yes
6	No	Medium	60K	No
7	Yes	Large	220K	No
8	No	Small	85K	Yes
9	No	Medium	75K	No
10	No	Small	90K	Yes

Eğitim Kümesi

Tid	Attrib1	Attrib2	Attrib3	Class
11	No	Small	55K	?
12	Yes	Medium	80K	?
13	Yes	Large	110K	?
14	No	Small	95K	?
15	No	Large	67K	?

Sınama Kümesi



- holdout
- repeated holdout
- k-fold cross validation
- bootstrapping



Büyük Veri Kümelerinde Değerlendirme

- Veri dağılımı dengeli ise: Veri kümesindeki örnek sayısı ve her sınıfa ait örnek sayısı fazla ise basit bir değerlendirme yeterli
 - holdout yöntemi: Belli sayıda örnek sınama için ayrılır, geriye kalan örnekler öğrenme için kullanılır
 - genelde veri kümesinin $\frac{2}{3}$ 'ü öğrenme, $\frac{1}{3}$ 'i sınama kümesi olarak ayrılır
 - öğrenme kümesi kullanılarak model oluşturulur ve sınama kümesi kullanılarak model değerlendirilir.



Tekrarlı Holdout Yöntemi

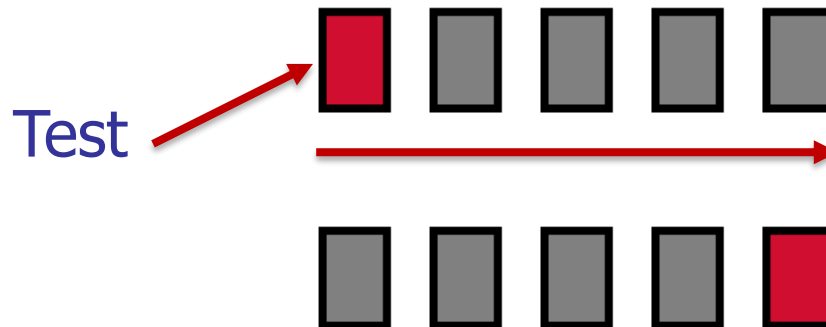
- Veri kümesini farklı altkümelere bölerek holdout yöntemini tekrarlama
 - Her eğitme işleminde veri kümesinin belli bir bölümü öğrenme kümesi olarak rasgele ayrılır.
 - Modelin hata oranı, işlemler sonunda elde edilen modellerin hata oranlarının ortalamasıdır.
- Problem: Farklı eğitme işlemlerindeki sınama kümeleri örtüşebilir.

k-Kat Çapraz Geçerleme

- Veri kümesi eşit boyutta k adet farklı gruba ayrılır.



- Bir grup sinama, diğerleri öğrenme için ayrılır.



- Her grup bir kere sinama kümesi olacak şekilde deneyler k kere tekrarlanır.

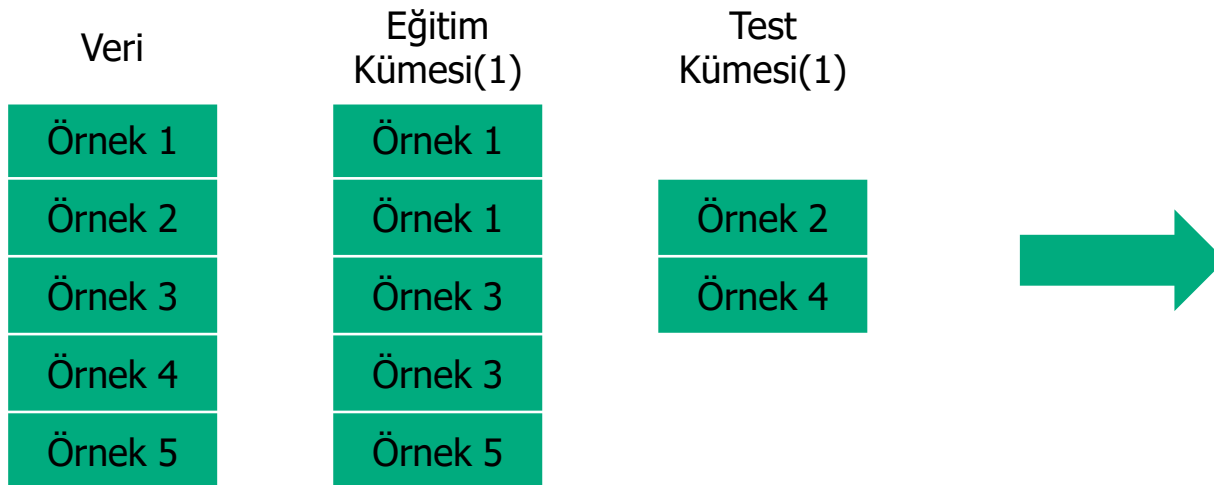


Biri Hariç Çapraz Geçerleme

- k-kat çapraz geçerlemenin özel hali
 - k sayısı veri kümesindeki örnek sayısına (N) eşit
- Model N-1 örnek üzerinde eğitilir, dışarıda bırakılan 1 örnek üzerinde sınanır.
- Bu işlem her örnek 1 kez sınama için kullanılacak şekilde tekrarlanır.
 - model N kez eğitilir.
- Model başarımı denemelerin başarımının ortalaması
- Verinin en etkin şekilde kullanımı

Bootstrap Yöntemi

- Veri kümesinden yerine koyma yöntemi ile örnekler seçilerek öğrenme kümesi oluşturulur.
 - N örnekten oluşan veri kümesinden yerine koyarak N örnek seçilir.
 - Bu küme öğrenme kümesi olarak kullanılır.
 - Öğrenme kümesinde yer almayan örnekler sınama kümesi olarak kullanılır.





0.632 bootstrap

- N örnekten oluşan bir veri kümesinde bir örneğin seçilmeme olasılığı: $1 - \frac{1}{N}$
- Sınama kümesinde yer alma olasılığı:
$$\left(1 - \frac{1}{N}\right)^N \approx e^{-1} = 0.368$$
- Öğrenme kümesi veri kümesindeki örneklerin %63,2'sinden oluşuyor.



Bootstrap Yönteminde Model Hatasını Belirleme

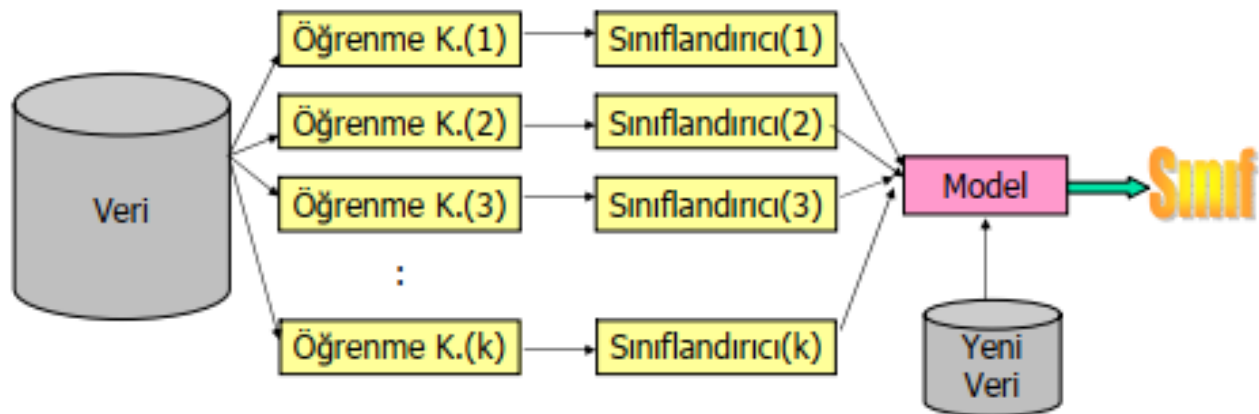
- Model başarımını sadece sinama kümesi kullanarak belirleme kötümser bir yaklaşım
 - model örneklerin sadece $\sim 63\%$ 'lük bölümüyle eğitiliyor
- Model başarımı hem öğrenme kümesindeki hem de sinama kümesindeki başarımla değerlendirilir.
$$\text{hata} = 0,632 \text{ hata}(\text{sinama}) + 0,368 \text{ hata}(\text{öğrenme})$$
- İşlem birkaç kez tekrarlanarak hatanın ortalaması alınır.



Sınıflandırıcıları Birleştirme

- Sınıflandırıcıları birleştirme
 - Bagging
 - Boosting

Model Başarımını Artırma



- Bir grup sınıflandırıcı kullanma
 - Bagging
 - Boosting



Bagging

- N örnekten oluşan bir veri kümesinde bootstrap yöntemi ile T örnek seç
- Bu işlemi k öğrenme kümesi oluşturmak üzere tekrarla
- Aynı sınıflandırma algoritmasını k öğrenme kümesi üzerinde kullanarak k adet sınıflandırıcı oluştur
- Yeni bir örneği sınıflandırmak için her sınıflandırıcının sonucunu öğren
- Yeni örnek en çok hangi sınıfa atanmışsa o sınıfın etiketiyle etiketlenir



Boosting

- Öğrenme kümesindeki her örneğin bir ağırlığı var.
- Her öğrenme işleminden sonra, her sınıflandırıcı için yapılan sınıflandırma hatasına bağlı olarak örneklerin ağırlığı güncelleniyor.
- Yeni bir örneği sınıflandırmak için her sınıflandırıcının doğruluğuna bağlı olarak ağırlıklı ortalaması alınıyor.