Makine Öğrenmesi

Metrikler, mesafe ölçüleri, sınıflandırma ölçütleri

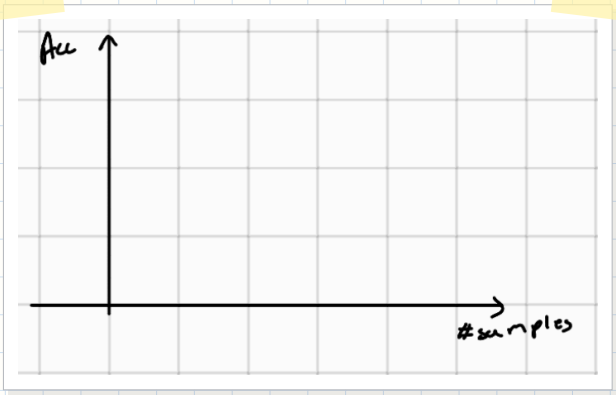
Orijinal not: Ders-2.pdf

UYARI Bu belge, el yazısı ders notlarından Claude (Anthropic) yardımıyla dijitale aktarılmıştır. Yardımcı kaynak olarak hazırlanmıştır; hata veya eksik içerebilir. Lütfen orijinal notları ve ders kaynaklarını esas alın.

NOTICE *This document was transcribed from handwritten lecture notes with the help of Claude (Anthropic). It is meant as a supplementary resource and may contain errors or omissions. Always check against the original notes and official course material.*

orj. s. 1

Independent component analysis



orijinal çizim, s. 1

zero shot mesela sınıfın yarısından örnek alındı, diğer tarafta test edildi. diğer taraftan örnek alınmamış.

Metrikler

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

Mesafe ölçümü ve metrikler:

Distance measure and metrics

Metrik:

① Negatif olmama: $d(x) \geq 0$[1]
② $d(x, y) = 0 \Rightarrow x = y$
③ $d(x, y) = d(y, x)$ simetri özelliği
④ $d(x, z) \leq d(x, y) + d(y, z)$

$\arg\max_x f(x)$, $f(x)$ max yapan

orj. s. 2

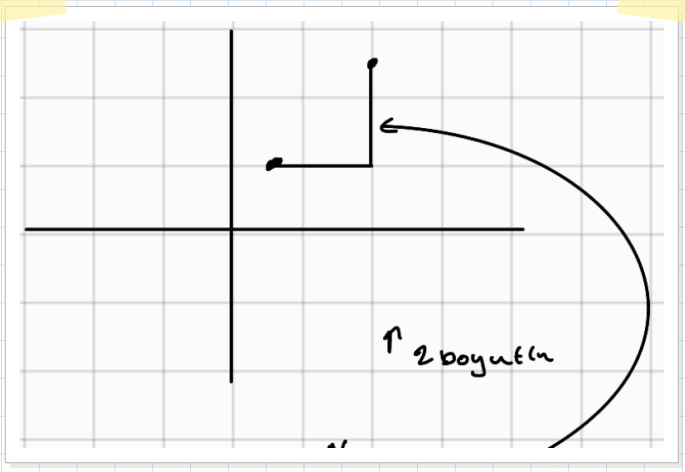
$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \qquad d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots}$$

Divergence: (Bu metrik değil ama o da mesafe ölçer) birisi diğerine daha uzak mı onu ölçüyor.

Tüm metrik özelliklerini sağlamazlar

norm:

$$\left.\begin{aligned} \|\vec{x}\|_2 &= \left(\sum_{i=1}^{N} |x_i|^2\right)^{1/2} \\ \|\vec{x}\|_p &= \left(\sum_{i=1}^{N} |x_i|^p\right)^{1/p} \end{aligned}\right\} \ell_p\text{-norm}$$

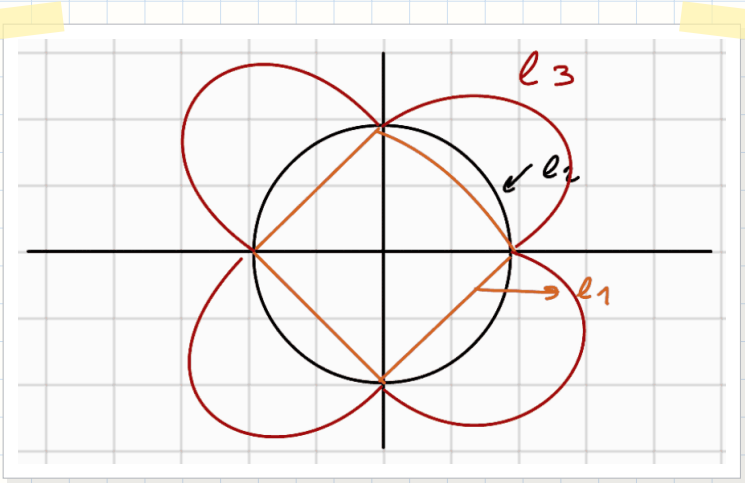


orijinal çizim, s. 2

$$\|\vec{x}\|_1 = \sum_{i=1}^{N} |x_i|$$

ℓ_1-norm: Manhattan norm
ℓ_2-norm: Öklid
ℓ_∞-norm: $\|x\|_\infty = \max_i |x_i|$

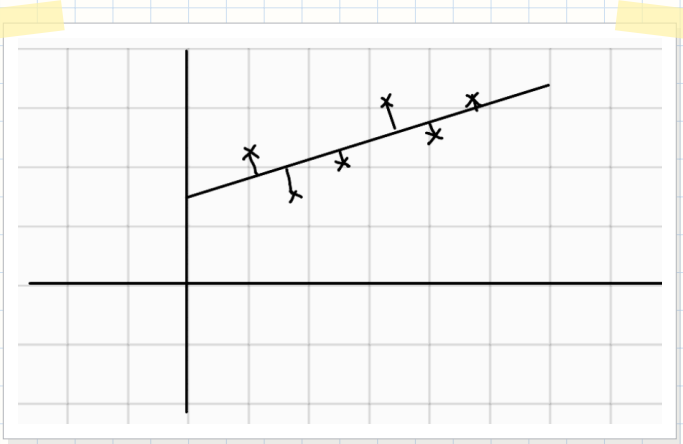
[1] Claude notu: muhtemelen $d(x, y) \geq 0$ olmalı.



orijinal çizim, s. 2

l_p-norma göre birim çemberler

orj. s. 3

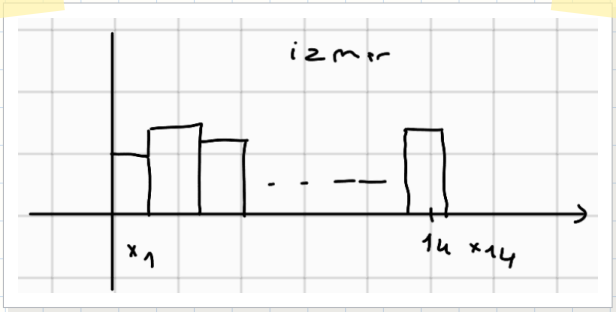


orijinal çizim, s. 3

→ Loss functionlarda normlar kullanılıyor

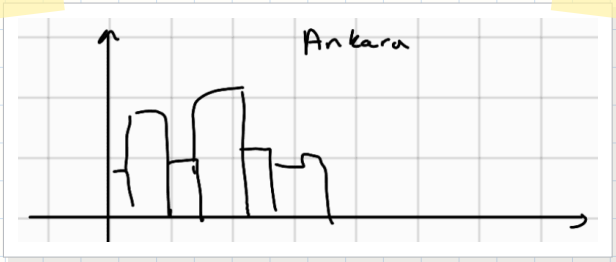
$$\|\alpha x\|_p = |\alpha| \, \|x\|_p$$

Okuma alışkanlığını şehirlere göre



orijinal çizim, s. 3

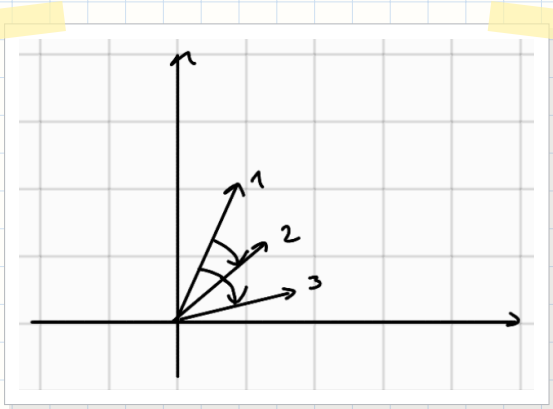
$$\vec{x}_{izmir} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{14} \end{bmatrix}$$



orijinal çizim, s. 3

burada sıkıntı ne
mesela nüfusa bölebiliriz

$$\|\vec{x}_{izmir} - \vec{x}_{ankara}\|_2 > \|\vec{x}_{izmir} - \vec{x}_{antep}\|_2$$



orijinal çizim, s. 3

$$\vec{x}_{izmir} \cdot \vec{x}_{ankara} = \frac{\|\vec{x}_{izmir}\| \, \|\vec{x}_{ankara}\| \cos\alpha}{\| \| \cdot \| \|}$$

orj. s. 4

Ama ben olasılık dağılımı üzerinden gitmek istiyorum

KL-Divergence:

$$D_{KL}(P\|Q) = \sum_{x \in \mathcal{X}} P(x) \cdot \log \frac{P(x)}{Q(x)}$$

bir kez izmir ankara
bir kez izmir antep

Ama bu metrik değil

çünkü

1) negatif olabilir
2) simetrik değil $d(x,y) \neq d(y,x)$ her zaman

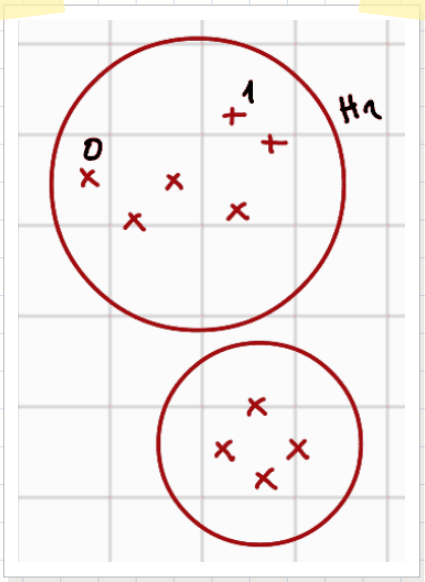
Metriğe çevirelim

$$D_{JS}(P,Q) = \begin{cases} \sum_{x\in\mathcal{X}} P(x) \left| \log \frac{P(x)}{Q(x)} \right| \\ + \sum_{x\in\mathcal{X}} P(x) \left| \log \frac{Q(x)}{P(x)} \right| \end{cases}$$

Jenson-Shannon div

$$M(x) = \frac{P(x) + Q(x)}{2}$$

$$\frac{1}{2} D_{KL}(P\|M) + \frac{1}{2} D_{KL}(Q\|M)$$



orijinal çizim, s. 4

Entropi:

$$H(x) = -\sum_{x\in\mathcal{X}} P(x) \log P(x)$$

$$P(0) \cdot \log P(0) + P(1) \log P(1)$$

$$H_1 = -\left(\frac{2}{3} \log \frac{2}{3} + \frac{1}{3} \log \frac{1}{3} \right)$$

$$H_2(x) = 0 + 0 = 0$$

Ne kadar çok kaos varsa entropi o kadar yükselir.

orj. s. 5 $P(0) = p \qquad P(1) = 1 - p \qquad p = 0{,}5$ max olur $H(x)$

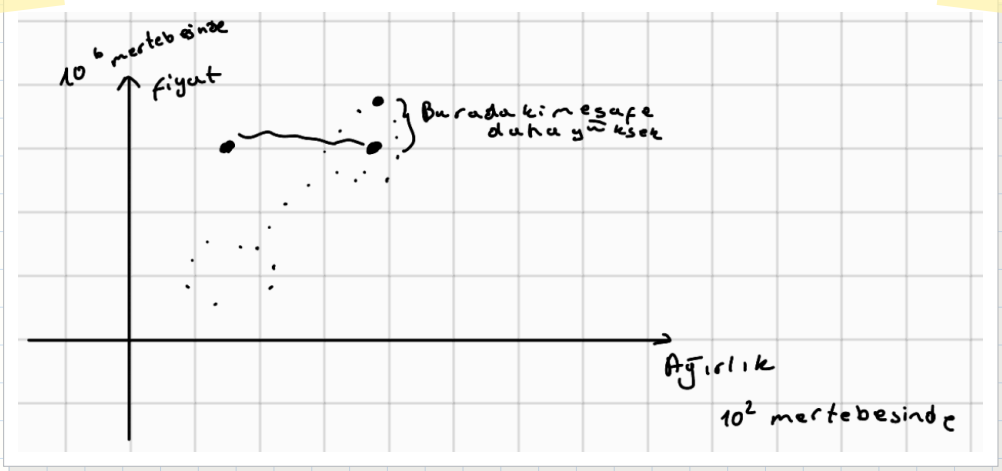
Continuos için:

$$H(x) = -\int_{-\infty}^{+\infty} p(x) \log p(x)\, dx$$

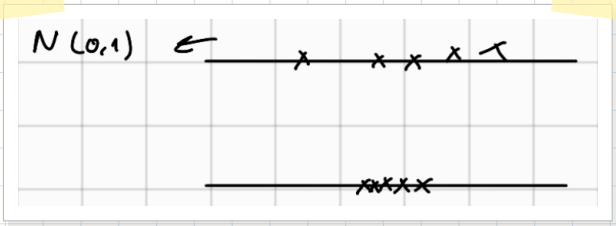
Cross-Entropy:

$$H(P, Q) = -\sum_{x \in \mathcal{X}} P(x) \log Q(x)$$

↑ P ile Q'nun birbirine uzaklığı

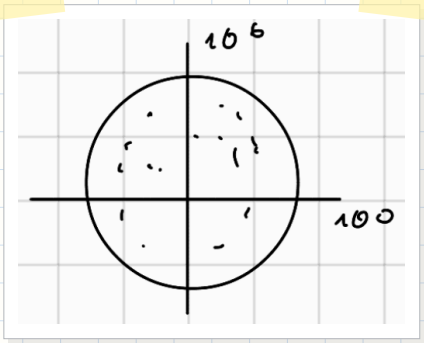


orijinal çizim, s. 5



orijinal çizim, s. 5

$$d_M(x, y) = \sqrt{(x-y)^T \underbrace{\Sigma_{xy}^{-1}}_{\text{kovaryans matrisinin tersi}} (x-y)}$$



orijinal çizim, s. 5

$$\Sigma = \begin{bmatrix} 10^{12} & 0 \\ 0 & 10^4 \end{bmatrix}$$

orj. s. 6

Bir haber başkasına benziyor mu

Iğdır'da trafik kazasından 10 kişi öldü

{ Tokenization

ığdır trafik kaza 10 kişi ölmek

gs fb 3 1 yemek

ığdır avokada getirmek

Küme Metrikleri

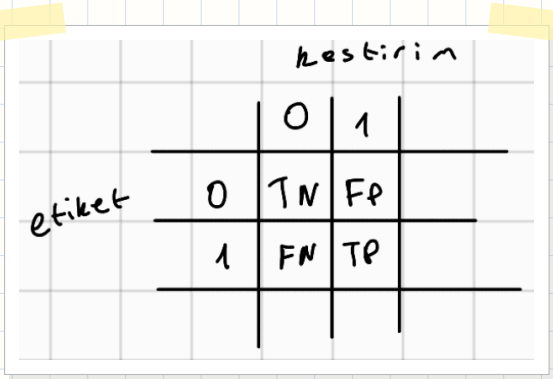
Hamming Distance Bütün veriler tek (set())

$$d_H(A,B) = |A \cap B|$$

Jaccard distance:

$$1 - \frac{|A \cap B|}{|A \cup B|}$$

Sınıflandırma metrikleri



orijinal çizim, s. 6

Doğruluk (Acc)

$$\frac{TN + TP}{\#total} = Acc$$

$$\{x^t, r^t\}_{t=1}^{N}$$

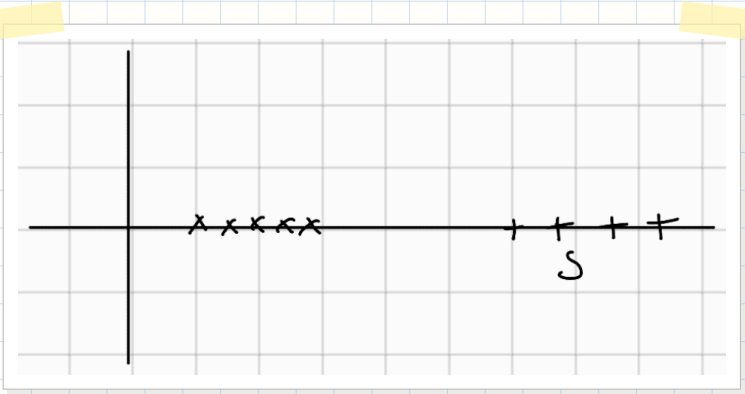
$$r^t = 0,\ t = 1{:}N$$

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

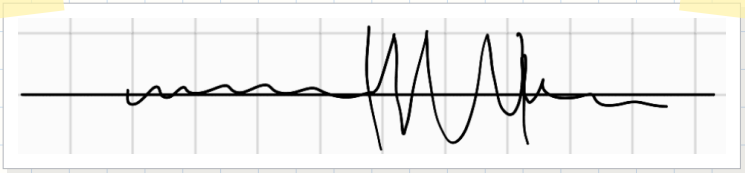
orj. s. 7

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

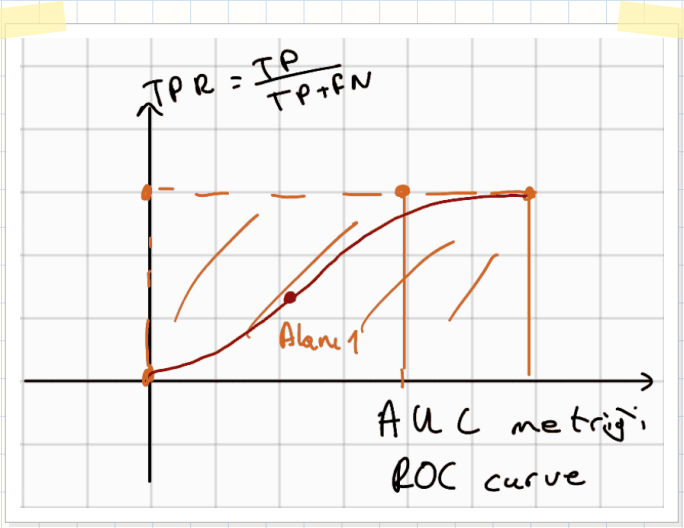
$$\text{F-1 score} = 2 \times \frac{\text{Prec} \times \text{Recell}}{\text{Prec} + \text{Recell}}$$



orijinal çizim, s. 7



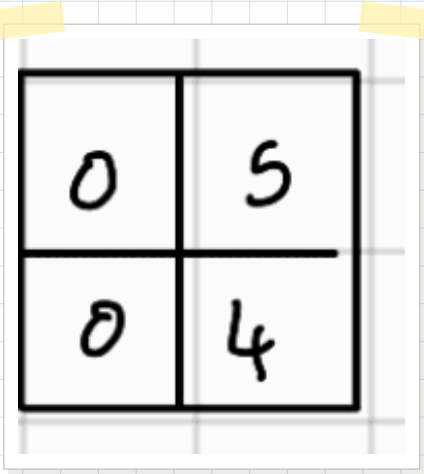
orijinal çizim, s. 7



orijinal çizim, s. 7

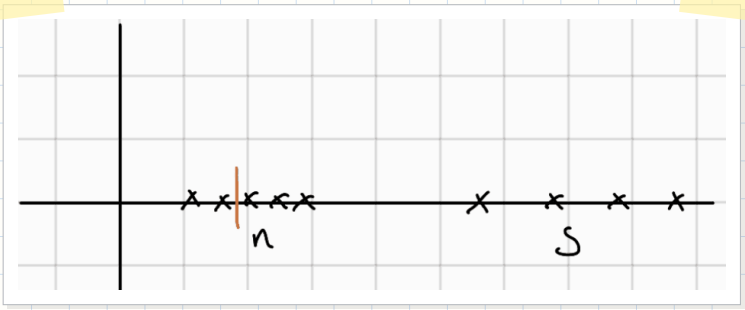
$$FPR = \frac{FP}{TN + FP}$$

0'ı eşik değer yaparsam

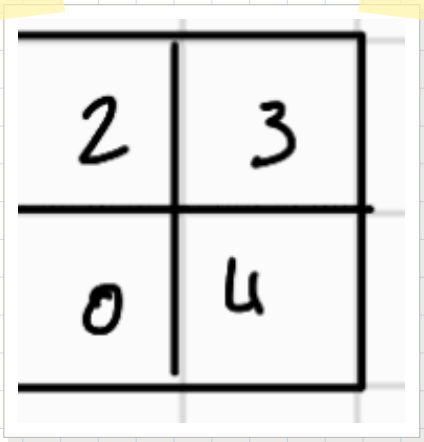


orijinal çizim, s. 7

$$TPR = \frac{4}{0+4} = 1 \qquad FPR = \frac{5}{5+0} = 1$$



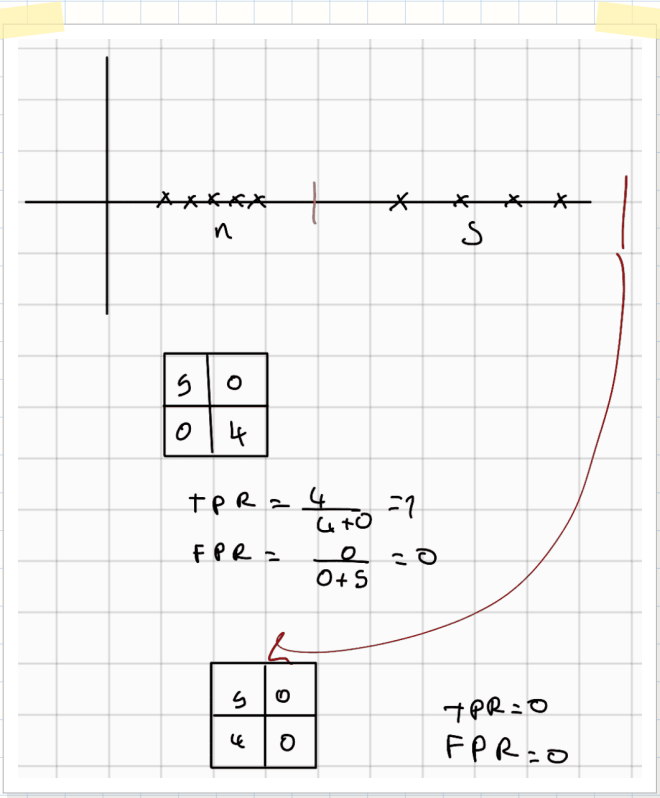
orijinal çizim, s. 7



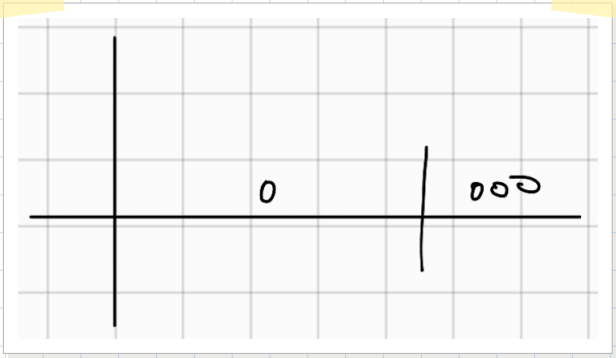
orijinal çizim, s. 7

$$TPR = \frac{4}{4+0} = 1 \qquad FPR = \frac{3}{3+2} = \frac{3}{5}$$

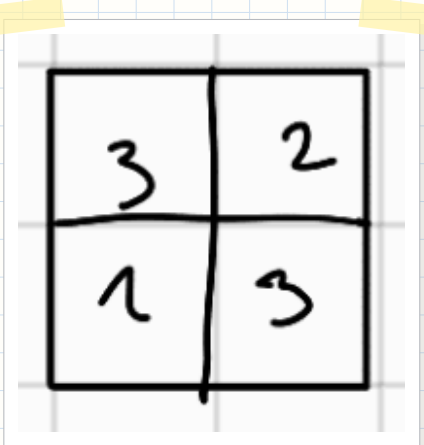
orj. s. 8 Soru



orijinal çizim, s. 8



orijinal çizim, s. 8

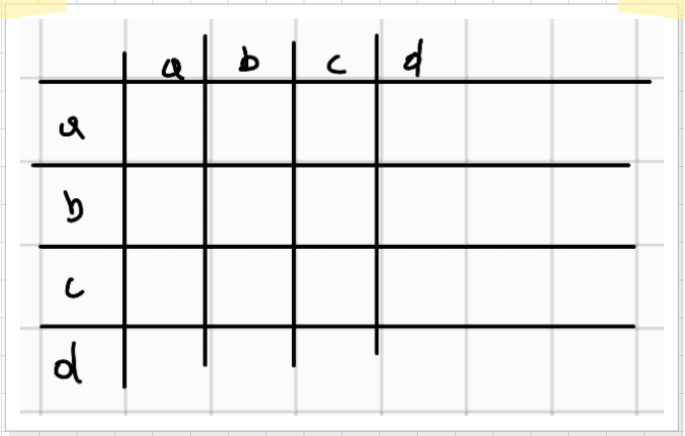


orijinal çizim, s. 8

$$TPR = \frac{2}{2+1} = \frac{2}{3} \qquad FPR = \frac{2}{2+3} = \frac{2}{5}$$

orj. s. 9

Confusion matrix

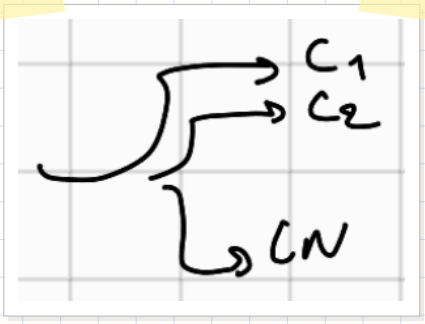


orijinal çizim, s. 9

$$\{x^t, r^t\}_{t=1}^N$$

$$r^t \in C$$

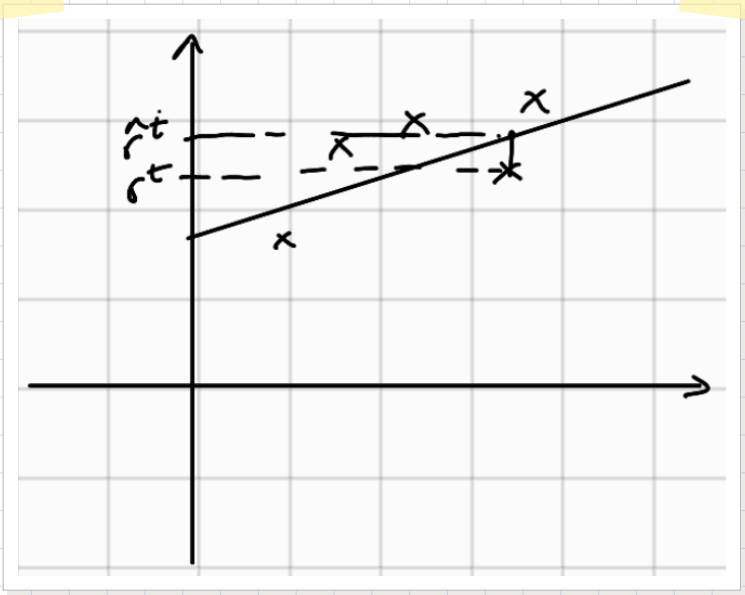
$$r^t \in \mathbb{R}^N$$



orijinal çizim, s. 9

$$f(x^t) = r^\alpha$$

$$\frac{1}{N}\sum_{t=1}^{N}\left(\underbrace{\hat{r}^t}_{\text{tahminler}} - r^t\right)^2$$



orijinal çizim, s. 9

$$ax + b = f(x)$$