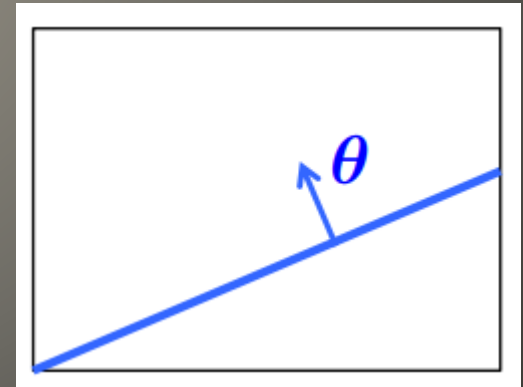


Doğrusal Sınıflandırma: Algılayıcı

Giriş
Doç. Dr. İlhan AYDIN

- Doğrusal Sınıflandırıcılar

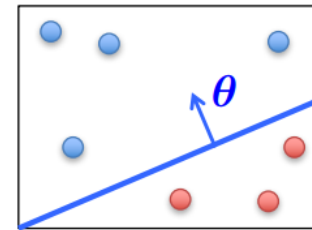
- Bir hiperdüzlem \mathbb{R}^d 'yi iki yarım uzaya böler
 - Normal vektör $\theta \in \mathbb{R}^d$ ile tanımlanır
 - θ , hiper düzlemde bulunan herhangi bir vektöre diktir.
- Orijinden geçtiği varsayılır
 - Bunun nedeni, θ_0 'e x_0 ile yanlılık terimini dahil etmemizdir.
- +1, -1 etiketleriyle sınıflandırmayı düşünün...



Doğrusal Sınıflandırma

- Doğrusal Sınıflandırıcılar
 - Doğrusal sınıflandırıcılar: karar sınırını hiper düzlemle temsil eder

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_d \end{bmatrix} \quad \mathbf{x}^\top = [1 \quad x_1 \quad \dots \quad x_d]$$



$$h(\mathbf{x}) = \text{sign}(\boldsymbol{\theta}^\top \mathbf{x}) \quad \text{where} \quad \text{sign}(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ -1 & \text{if } z < 0 \end{cases}$$

– Note that: $\boldsymbol{\theta}^\top \mathbf{x} > 0 \implies y = +1$

$\boldsymbol{\theta}^\top \mathbf{x} < 0 \implies y = -1$

Doğrusal Sınıflandırma

- Algılayıcı

$$h(\mathbf{x}) = \text{sign}(\boldsymbol{\theta}^\top \mathbf{x}) \quad \text{where} \quad \text{sign}(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ -1 & \text{if } z < 0 \end{cases}$$

- Algılayıcı, her yeni eğitim örneği aldığı anda aşağıdaki güncelleme kuralını kullanır (x^i, y^i)

$$\theta_j \leftarrow \theta_j - \frac{\alpha}{2} \left(\underbrace{h_{\boldsymbol{\theta}}(\mathbf{x}^{(i)}) - y^{(i)}}_{\text{either 2 or -2}} \right) x_j^{(i)}$$

- Tahmin etiketle eşleşirse değişiklik yapmayın
- Aksi takdirde, θ olarak ayarlayın

Doğrusal Sınıflandırma

- Algılayıcı

$$h(\mathbf{x}) = \text{sign}(\boldsymbol{\theta}^\top \mathbf{x}) \quad \text{where} \quad \text{sign}(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ -1 & \text{if } z < 0 \end{cases}$$

- Algılayıcı, her yeni eğitim örneği aldığı anda aşağıdaki güncelleme kuralını kullanır (x^i, y^i)

$$\theta_j \leftarrow \theta_j - \frac{\alpha}{2} \underbrace{\left(h_{\boldsymbol{\theta}}(\mathbf{x}^{(i)}) - y^{(i)} \right)}_{\text{either 2 or -2}} x_j^{(i)}$$

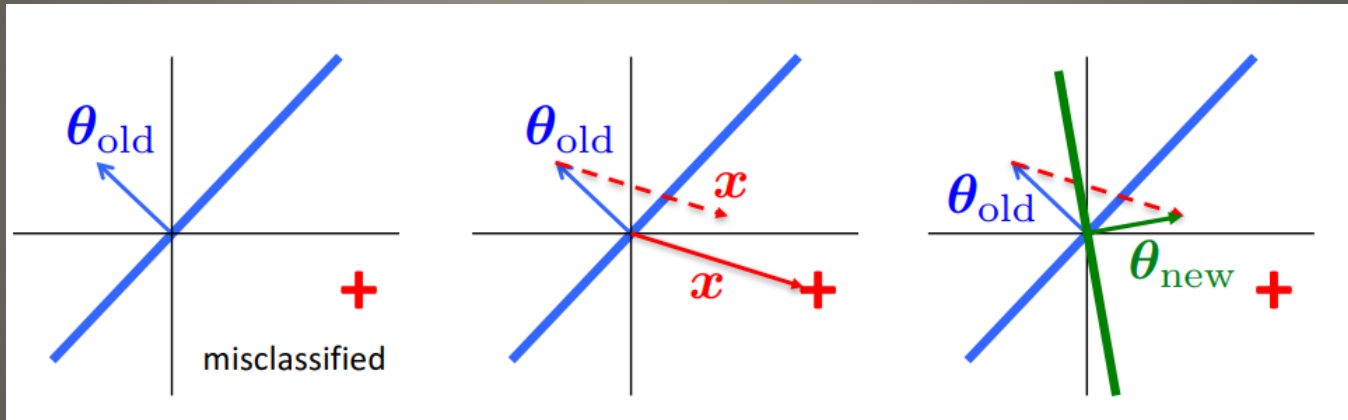
- Olarak yazın (sadece yanlış sınıflandırma üzerine)

$$\theta_j \leftarrow \theta_j + \alpha y^{(i)} x_j^{(i)}$$

- Bu durumda α 'yı ortadan kaldırabilir, çünkü tek etkisi θ 'yi performansı etkilemeyen bir sabitle ölçeklendirmektir.

Doğrusal Sınıflandırma

- Algılayıcı Güncellemesi Neden Çalışır?



Doğrusal Sınıflandırma

- Algılayıcı Güncellemesi Neden Çalışır?
 - Yanlış sınıflandırılmış örneği düşünün ($y = +1$)
 - Algılayıcı yanlış bir şekilde $\theta_{old}^T x < 0$ olduğunu düşünür
 - Güncelleme: $\theta_{new} = \theta_{old} + yx = \theta_{old} + x$ ($y = +1$ olduğundan)
 - Not: $\theta_{new}^T x = (\theta_{old} + x)^T x$
 - Bu nedenle, $\theta_{new}^T x$, θ_{old}^T 'den daha az negatiftir
 - Yani, bu örnekte kendimizi daha doğru yapıyoruz!

Doğrusal Sınıflandırma

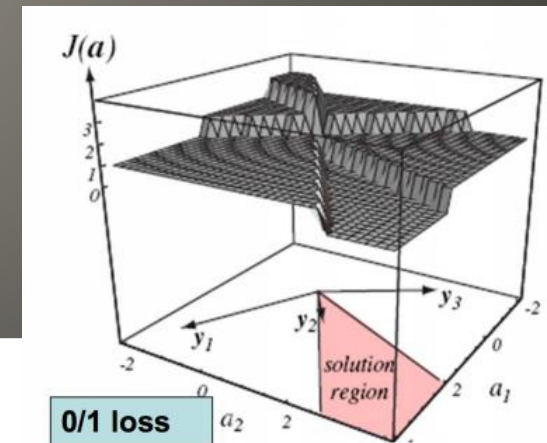
- Algılayıcı Maliyet Fonksiyonu

- Eğer $y^i \theta^T x^i > 0$ ise tahmin doğrudur.
- 0/1 kayıp kullanabilirdi

$$J_{0/1}(\theta) = \frac{1}{n} \sum_{i=1}^n \ell(\text{sign}(\theta^T x^{(i)}), y^{(i)})$$

- Burada $\ell()$, tahmin doğruysa 0, aksi takdirde 1'dir.
- Kullanışlı bir gradyan oluşturmaz

Doğrusal Sınıflandırma

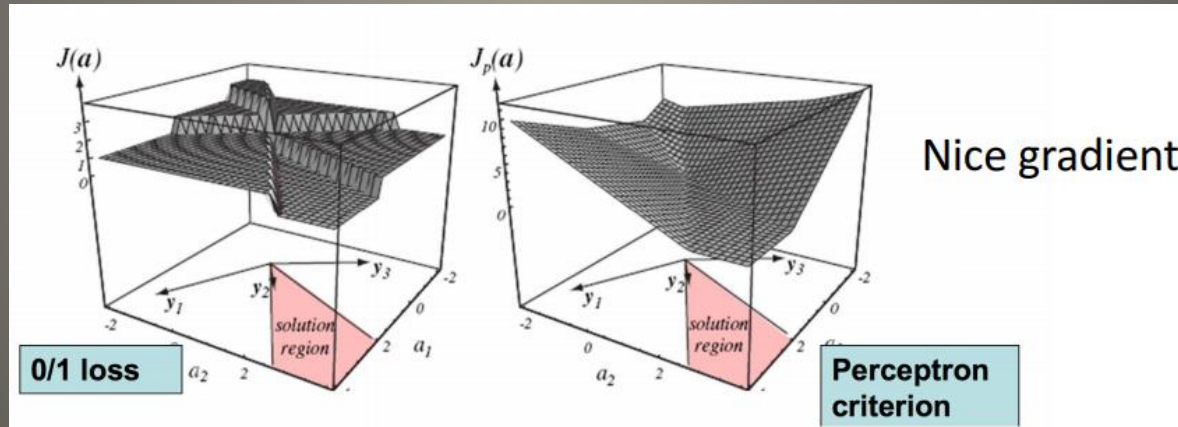


- Algılayıcı Maliyet Fonksiyonu

- Algılayıcı" aşağıdaki maliyet fonksiyonunun kullanılır.

$$J_p(\theta) = \frac{1}{n} \sum_{i=1}^n \max(0, -y^{(i)} \theta^T x^{(i)})$$

- Tahmin doğruysa $\max(0, -y\theta^T x^i)$ 0'dır
- Aksi takdirde, yanlış tahmine olan güvendir.



Doğrusal Sınıflandırma

- Çevrimiçi Algılayıcı Algoritması

Let $\theta \leftarrow [0, 0, \dots, 0]$

Repeat:

Receive training example $(\mathbf{x}^{(i)}, y^{(i)})$

if $y^{(i)} \mathbf{x}^{(i)} \theta \leq 0$

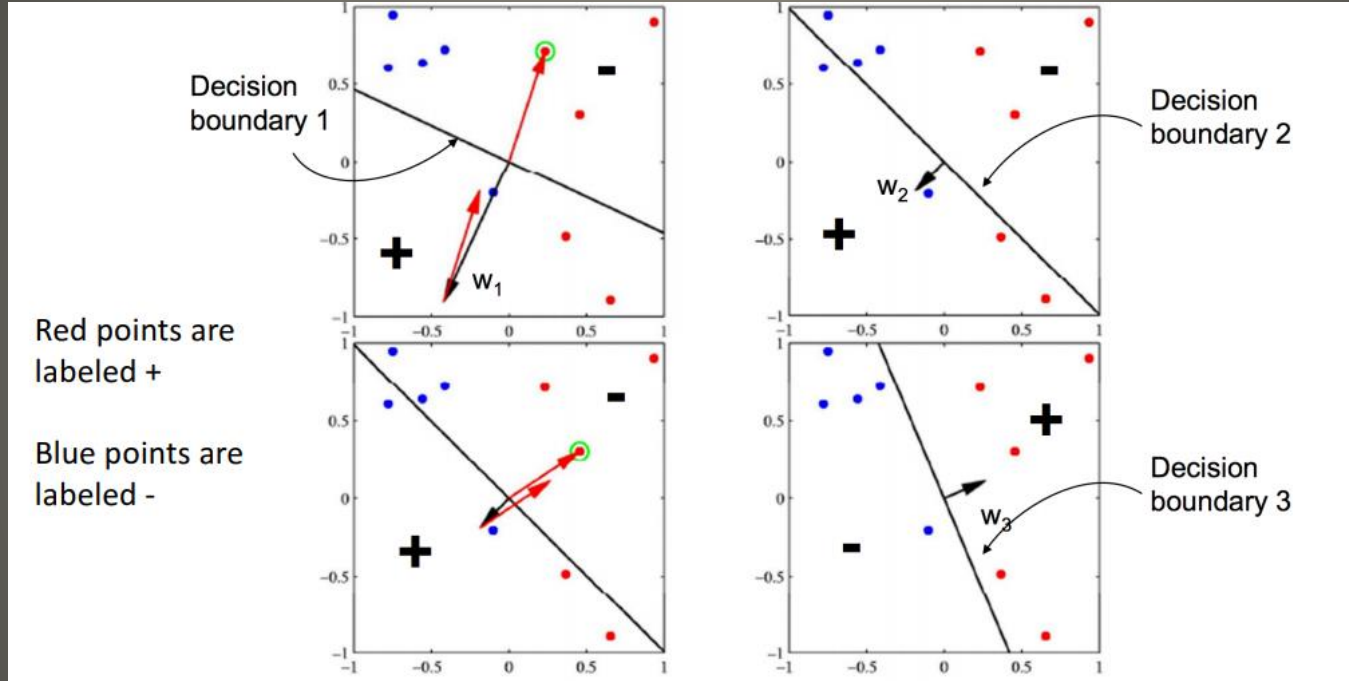
// prediction is incorrect

$\theta \leftarrow \theta + y^{(i)} \mathbf{x}^{(i)}$

- Çevrimiçi (Online) öğrenme – her tek gözlem alındığında model güncellemesinin gerçekleştirildiği öğrenme modu
- Toplu (Batch) öğrenme – tüm eğitim setini gözlemledikten sonra model güncellemesinin gerçekleştirildiği öğrenme modu

Doğrusal Sınıflandırma

- Çevrimiçi Algılayıcı Algoritması
 - Bir hata yapıldığında, ağırlığı hatayı düzelten yönde hareket ettirir.



- Algılayıcıyı çalışırken görün:
www.youtube.com/watch?v=vGwemZhPlsA

Doğrusal Sınıflandırma

- Toplu Algılayıcı (Batch Perceptron)

```
Given training data  $\{(\mathbf{x}^{(i)}, y^{(i)})\}_{i=1}^n$   
Let  $\boldsymbol{\theta} \leftarrow [0, 0, \dots, 0]$   
Repeat:  
    Let  $\boldsymbol{\Delta} \leftarrow [0, 0, \dots, 0]$   
    for  $i = 1 \dots n$ , do  
        if  $y^{(i)} \mathbf{x}^{(i)} \boldsymbol{\theta} \leq 0$  // prediction for  $i^{th}$  instance is incorrect  
             $\boldsymbol{\Delta} \leftarrow \boldsymbol{\Delta} + y^{(i)} \mathbf{x}^{(i)}$   
     $\boldsymbol{\Delta} \leftarrow \boldsymbol{\Delta} / n$  // compute average update  
     $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + \alpha \boldsymbol{\Delta}$   
Until  $\|\boldsymbol{\Delta}\|_2 < \epsilon$ 
```

- En basit durum: $\alpha = 1$ ve normalleştirmeyin, sabit artış algılayıcısını verir
- Varsa, ayırıcı bir hiper düzlem bulma garantisi

Doğrusal Sınıflandırma

- Algılayıcıyı Geliştirmek

- Algılayıcı, eğitim sırasında birçok θ üretir
- Standart Algılayıcı, test zamanında sadece son θ 'yi kullanır
 - Bu bazen iyi bir fikir olmayabilir!
 - Diğer bazı θ ardışık 1000 örnekte doğru olabilir, ancak bir hata onu mahveder!
- Fikir: Birden çok algılayıcının bir kombinasyonunu kullanın
 - (Örneğin, sinir ağları!)
- Fikir: Ara θ 'leri kullanın
 - Oylanan (Voted) Algılayıcı: ara θ 'lerin tahminlerine oy verin
 - Ortalama (Averaged) Algılayıcı: ara θ 'lerin ortalaması

Doğrusal Sınıflandırma