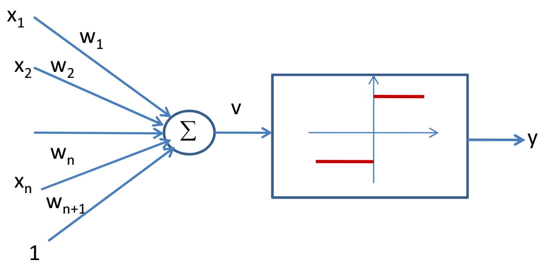
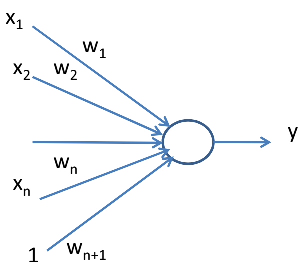
Hatırlatma

Genlikte Ayrık Algılayıcı-GAA (Perceptron)





$$v = \begin{bmatrix} w_1 & w_2 & \dots & w_n & w_{n+1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

$$v = w_1x_1 + w_2x_2 + \ldots + w_nx_n + w_{n+1}1$$

$$y = \varphi(v) = \begin{cases} 1 & v \geq 0 \\ -1 & v < 0 \end{cases}$$

Hatırlatma

Öğrenme Kuralı:

$$x \in S_1 \qquad w^T x > 0 \qquad w(k+1) = w(k)$$

$$x \in S_2 \qquad w^T x \leq 0 \qquad w(k+1) = w(k)$$

$$x \in S_1 \qquad w^T x \leq 0 \qquad w(k+1) = w(k) + cx$$

$$x \in S_2 \qquad w^T x > 0 \qquad w(k+1) = w(k) - cx$$

öğrenme hızı <1 olan pozitif bir sayı

$$\Delta w_i = c\frac{1}{2}[y_d - y]x_i$$

$$\Delta w = c\frac{1}{2}[y_d - y]x$$

Genlikte Ayrık Algılayıcı Yakınsama Teoremi

Eğitim kümesi doğrusal ayrıştırılabilir ise genlikte ayrık algılayıcı için verilen öğrenme kuralı sonlu adımda bir çözüm verir.

Tanıt: c=1

X eğitim kümesindeki tüm girişler olsun

$$x^{(1)}, x^{(2)}, \ldots, x^{(k)}, \ldots$$

Bu eğitim kümesine karşı düşen ağırlık vektörleri

$$w^{(1)}, w^{(2)}, \ldots, w^{(k)}, \ldots$$

$w^{(k+1)} = w^{(k)}$ olan adımlarda eğitim kümesindeki örüntüler ve ağırlık vektörleri yukarıdaki dizilerden çıkarılsın

Geriye hangileri kaldı?

Ağırlıkların güncellenmesi gereken her adımda ne oluyor?

$$w^{(j)T} x^{(j)} \leq 0 \ \text{ve} \ x^{(j)} \in S_1$$

$$w^{(j)T} x^{(j)} > 0 \ \text{ve} \ x^{(j)} \in S_2$$

Güncelleme sırasında neler olacak

$$w^{(2)} = w^{(1)} + x^{(1)}$$

$$w^{(3)} = w^{(2)} + x^{(2)} = w^{(1)} + x^{(1)} + x^{(2)}$$

$$w^{(4)} = w^{(3)} + x^{(3)} = w^{(1)} + x^{(1)} + x^{(2)} + x^{(3)}$$

Burada tuhaf olan birşey var.

$$\vdots$$

$$w^{(k+1)} = w^{(k)} + x^{(k)} = w^{(1)} + x^{(1)} + \ldots + x^{(k)}$$

öğrenme kuralı

$$x \in S_1 \qquad w^T x \leq 0 \qquad w(k+1) = w(k) + x$$

$$x \in S_2 \qquad w^T x > 0 \qquad w(k+1) = w(k) - x$$

Eğitim kümesi doğrusal ayrıştırılabilir

$\rightarrow$ $\exists\,\Omega$ çözüm kümesi $\ni\ w\in\Omega,\ \ w^T x>0$

Burada da tuhaf olan
birşey var.

$\rightarrow$ Bir $w^*\in\Omega$ seçilsin $\quad \alpha \mathrel{\hat=} \min\limits_{x\in X} w^{*T}x$

$$\beta \mathrel{\hat=} w^{*T} w^{(1)}$$

$$w^{*T}w^{(k+1)} = \underbrace{w^{*T}w^{(1)}}_{\beta} + \underbrace{w^{*T}x^{(1)}}_{\ge\alpha} + \dots . + \underbrace{w^{*T}x^{(k)}}_{\ge\alpha}$$

$$w^{*T}w^{(k+1)} \ge \beta + k\alpha$$

$$\left\|w^{*T}\right\|^2\left\|w^{(k+1)}\right\|^2 \ge \left\|w^{*T}w^{(k+1)}\right\|^2 \ge (\beta+k\alpha)^2$$

Alt sınır

$$\left\|w^{(k+1)}\right\|^2 \ge \frac{(k\alpha+\beta)^2}{\left\|w^*\right\|^2}$$

$$w^{(i+1)} = w^{(i)} + x^{(i)}$$

$$\left\|w^{(i+1)}\right\|^2 = w^{(i+1)T}w^{(i+1)} = w^{(i)T}w^{(i)} + 2w^{(i)T}x^{(i)} + x^{(i)T}x^{(i)}$$

$$= \left\|w^{(i)}\right\|^2 + 2w^{(i)T}x^{(i)} + \left\|x^{(i)}\right\|^2$$

Neden?

$w^{(i)T}x^{(i)} \le 0 \ \rightarrow\ \left\|w^{(i+1)}\right\|^2 - \left\|w^{(i)}\right\|^2 \le \left\|x^{(i)}\right\|^2$

Nasıl geldik?

$$\forall i = 1,2,3,\dots,k$$

$A \mathrel{\hat=} \max\limits_{x\in X}\|x\|^2 \ \rightarrow\ \left\|w^{(k+1)}\right\|^2 \le \sum_{i=1}^{k}\left\|x^{(i)}\right\|^2 + \left\|w^{(1)}\right\|^2$

$$\le kA + \left\|w^{(1)}\right\|^2$$

Üst sınır

$$\left\|w^{(k+1)}\right\|^2 \le kA + \left\|w^{(1)}\right\|^2$$

$$\left\|w^{(k+1)}\right\|^2 \ge \frac{(k\alpha+\beta)^2}{\left\|w^*\right\|^2} \qquad \left\|w^{(k+1)}\right\|^2 \le kA + \left\|w^{(1)}\right\|^2$$

Alt sınır

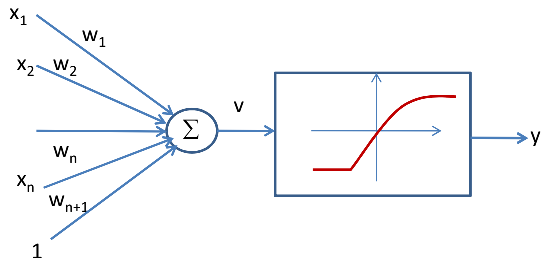
Büyük k'lar için bu
sınırlara dikkat edin!!!!

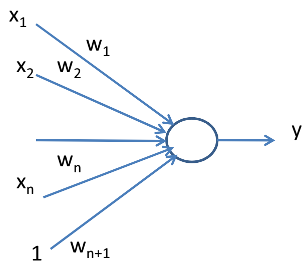
Üst sınır

$$\frac{(k_m\alpha+\beta)^2}{\left\|w^*\right\|^2} = k_m A + \left\|w^{(1)}\right\|^2$$

en büyük adım sayısı yukarıdaki eşitliğin belirlediği k_m 'den büyük olamaz ■

Genlikte Sürekli Algılayıcı (ADALİNE)





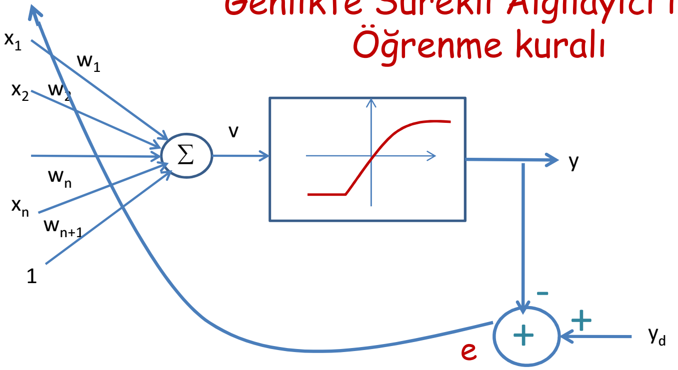
$$v = \begin{bmatrix} w_1 & w_2 & \dots . & w_n & w_{n+1}\end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$$

$$v = w_1x_1 + w_2x_2 + \dots + w_nx_n + w_{n+1}1$$

$$y = \varphi(v) = \tanh av$$

$$y = \varphi(v) = \frac{1}{1+e^{-av}}$$

Genlikte Sürekli Algılayıcı için Öğrenme kuralı



Amaç: hatayı azaltacak ağırlıkları belirlemek

Hataya ilişkin bir fonksiyon oluşturularak işe başlanacak

$$e = y_d - y$$

Neden bir fonksiyon?

$$E = \frac{1}{2} e^T e = (y_d - y)^T (y_d - y) = (y_d - \varphi(v))^T (y_d - \varphi(v))$$

Nasıl azaltabiliriz?

Amaç: $\min_{w \in R^{n+1}} E$

Neden vazgeçtik?

$$E = \frac{1}{2} e^T e = (y_d - y)^T (y_d - y) = (y_d - \varphi(v))^T (y_d - \varphi(v))$$
$$= (y_d - \varphi(w_1 x_1 + w_2 x_2 + \ldots + w_n x_n + w_{n+1} 1))^T$$
$$(y_d - \varphi(w_1 x_1 + w_2 x_2 + \ldots + w_n x_n + w_{n+1} 1))$$

Böylece E'nin w'ya bağımlılığını açıkça yazdık, acaba $\min_{w \in R^{n+1}} E$ sağlayan w'ları nasıl buluruz?

$E(w^{(k+1)}) \le E(w^{(k)})$ sağlayacak $w^{(k+1)}$'i $w^{(k)}$ 'dan nasıl elde ederiz?

$$w^{(k)} + ? = w^{(k+1)} \quad \Rightarrow \quad E(w^{(k+1)}) \le E(w^{(k)})$$

$$w^{(k)} - c\nabla E = w^{(k+1)}$$

Nasıl bulunur?

$$\nabla E \,\hat{=} \begin{bmatrix} \frac{\partial E}{\partial w_1} \\ \frac{\partial E}{\partial w_2} \\ \bullet \\ \bullet \\ \bullet \\ \frac{\partial E}{\partial w_{n+1}} \end{bmatrix} = \frac{\partial E}{\partial e}\frac{\partial e}{\partial y}\frac{\partial y}{\partial v}\frac{\partial v}{\partial w} = -(y_d - y)\varphi'(v) \begin{bmatrix} \frac{\partial v}{\partial w_1} \\ \frac{\partial v}{\partial w_2} \\ \bullet \\ \bullet \\ \bullet \\ \frac{\partial v}{\partial w_{n+1}} \end{bmatrix}$$

$$= -(y_d - y)\varphi'(v) \begin{bmatrix} x_1 \\ x_2 \\ \bullet \\ \bullet \\ \bullet \\ 1 \end{bmatrix}$$

Öğrenme Kuralı:

$$\Delta w = -(y_d - y)\varphi'(v) x$$

Aktivasyon fonksiyonunu lineer alırsak ...

$$\Delta w = -(y_d - y)\varphi'(v) x \quad \Rightarrow \quad \Delta w = -(y_d - y) x$$

Bir de girişleri normalize edersek ...

$$\Delta w = -(y_d - y)\varphi'(v) x \quad \Rightarrow \quad \Delta w = -(y_d - y)\frac{x}{\|x\|}$$

ADALİNE'ı nerede kullanabiliriz

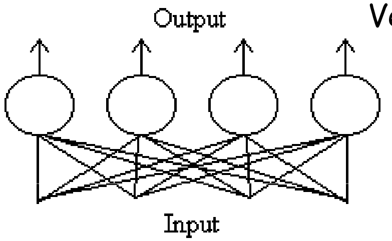
- sınıflandırma probleminde

Eğitim aşamasından sonra $y = \begin{cases} 1 & \sum_{i=1}^{n} w_{ij} x_i > 0 \\ -1 & \sum_{i=1}^{n} w_{ij} x_i \le 0 \end{cases}$

- yaklaşık eğri uydurma
- Lineer regresyon

Hep iki sınıfa ayırdık daha fazla sınıfa nasıl ayıracağız?

Unutulmaması gereken kısıt ne?



http://richardbowles.tripod.com/neural/slpmlp/

Verilenler: $\left\{\left(x^l, y_d^l\right)\right\}\Big|_{l=1}^{P}$ Eğitim Kümesi

Amaç: m sınıfa ayırmak

$$S_1 : \sum_{i=1}^{n} w_i x_i + w_{n+1} < T_1$$

$$S_2 : T_1 \leq \sum_{i=1}^{n} w_i x_i + w_{n+1} < T_2$$

.....

$$S_2 : T_{m-1} \leq \sum_{i=1}^{n} w_i x_i + w_{n+1} < T_m$$

Öğrenme Kuralı: $\Delta w = c\frac{1}{2}[y_d - y]x$ $\quad \Delta w = -(y_d - y)\varphi'(v)x$