SUPPORT VECTOR MASCHINEN (SVM) · Vereinfachte Form.

Das Ziel von SVM ist die optimale Trennlinie zu finden, welche vorgegebene Trainingsdaten (bereits gelabelt) trennt.

Beispiel. Gegeben sind folgende Trainingsdaten

ROTE (A): [1,1], [2,3], [3,2] ; BLAUE (B): [5,3], [6,2], [6,4]

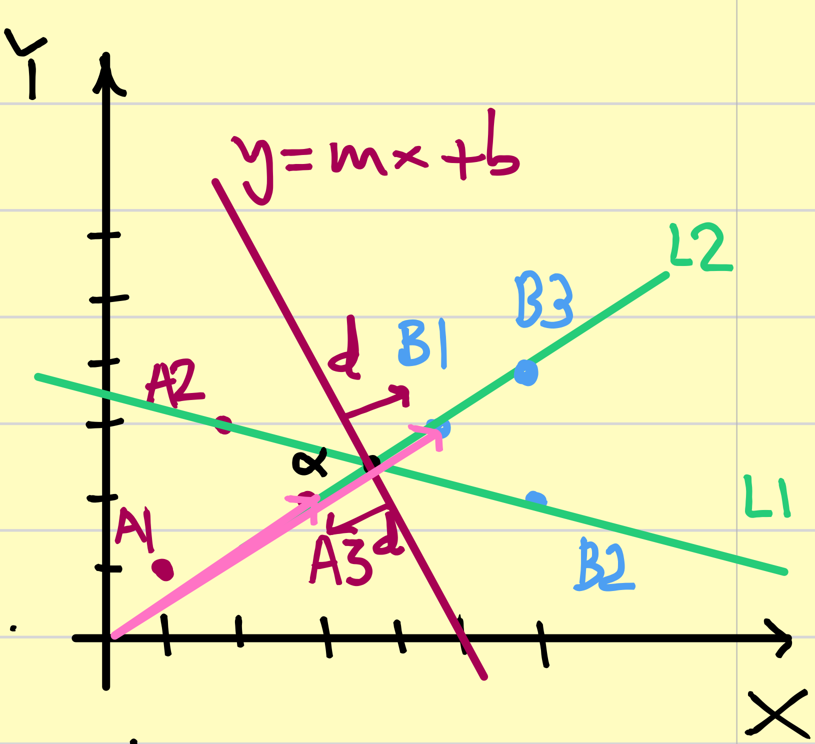
Ziel: a) finden Sie die optimale Trennlinie ($y = mx + b$), welche die Klassen am besten separiert.

b) berechnen Sie die Abstände (d). MARGIN. der nächstgelegenen Punkten zur Trennlinie.

c) definieren Sie die Support Vektoren.

SCHRITT 1. Punkte visualisieren.

SCHRITT 2. Trennlinie in der Form $y = mx + b$. Der Abstand der Trennlinie zu den nächstgelegen Punkten der beiden Klassen sollte maximiert werden und gleich sein.



Der Abstand zw. eine Linie $y = mx + b$ und einen Punkt mit Koord. $[x_1, y_1]$ lautet:

$$d = \frac{|mx_1 + y_1 + b|}{\sqrt{m^2+1}}$$

SCHRITT 3: Hilfe linien, welche die Daten trennen.

Die linien L1 & L2 trennen die Daten und da wo sie sich kreuzen [PUNKT α] geht unsere optimale Trennlinie.

L1: A2 , B2 :
[2,3] [6,2]

$$\boxed{\frac{y-y_1}{x-x_1} = \frac{y_2-y_1}{x_2-x_1}}$$

Eine linie die durch $[x_1, y_1]$ und $[x_2, y_2]$ geht.

$$\frac{y-2}{x-6} = \frac{3-2}{2-6} \longrightarrow y-2 = \frac{-1}{4}(x-6) \qquad (1)$$

L2: A3 , B3
[3,2] [6,4]

$$\frac{y-4}{x-6} = \frac{2-4}{3-6} \longrightarrow y-4 = \frac{2}{3}(x-6) \qquad (2)$$

$$(1)-(2) \longrightarrow (\cancel{y_\alpha}-2)-(\cancel{y_\alpha}-4) = \frac{-1}{4}(x_\alpha-6) - \frac{2}{3}(x_\alpha-6)$$

$$2 = \left(\frac{-1}{4} - \frac{2}{3}\right)(x_\alpha-6) \longrightarrow 2 = \frac{-11}{12}(x_\alpha-6)$$

$$\longrightarrow x_\alpha = 3'818$$

$$(1) \qquad y_\alpha = 2 - \frac{1}{4}(3'818-6) = 2'545$$

$$\alpha = [3'818, 2'545]$$

SCHRITT 4.

d: B1 : $d = \dfrac{|5m+3+b|}{\sqrt{m^2+1}}$ (3)

[5,3]

d: A3 : $d = \dfrac{|3m+2+b|}{\sqrt{m^2+1}}$ (4)

[3,2]

Trennlinie geht durch α : $y_\alpha = m x_\alpha + b \rightarrow 2{,}545 = 3{,}818m + b$ (5)

Mit diesen drei Gleichungen, können wir m, b & d ermitteln

$(3)=(4) \rightarrow 5m+3+\cancel{b} = 3m+2+\cancel{b} \rightarrow m = \frac{-1}{2}$

$(5) \rightarrow 2{,}545 = 3{,}818 \cdot \left(\frac{-1}{2}\right) + b \rightarrow b = 0{,}636$

$(3) \rightarrow d = \dfrac{\left|5 \cdot \frac{-1}{2} + 3 + 0{,}636\right|}{\sqrt{\left(\frac{-1}{2}\right)^2+1}} = 1{,}016$

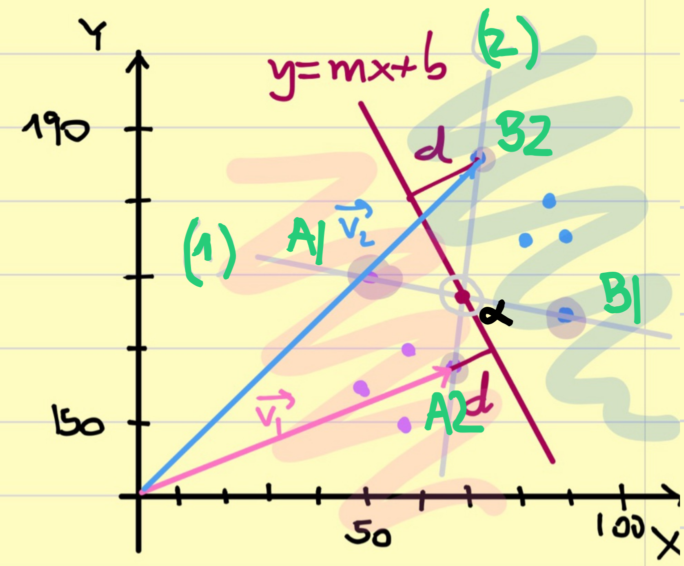
$\boxed{y = \frac{-1}{2}x + 0{,}636}$ $\boxed{d = 1{,}016}$

Support Vektoren A1 & B3

Beispiel 2.

Beispiel

	x (Gewicht)	y (Größe)	Klasse
	50	155	♀
	60	160	♀
A2	68	158	♀
	58	150	♀
A1	55	170	♀
B1	90	165	♂
	85	175	♂
	88	180	♂
B2	75	190	♂
	72	185	♂



(1): A1 B1
[55,170] [90,165]

$$\frac{y-170}{x-55} = \frac{165-170}{90-55} \rightarrow y-170 = -0'143(x-55) \quad (1)$$

(2): A2 B2
[68,158] [75,190]

$$\frac{y-158}{x-68} = \frac{190-158}{75-68} \rightarrow y-158 = 4'57(x-68) \quad (2)$$

$$(1)-(2): \; (\cancel{y_\alpha}-170)-(\cancel{y_\alpha}-158) = -0'143(x_\alpha-55) - 4'57(x_\alpha-68)$$

$$-12 = -0'143\,x_\alpha + 7'685 - 4'57\,x_\alpha + 310'76$$

$$x_\alpha = \frac{330'445}{4'713} = 70'11$$

$$(1): \; y_\alpha = 170 - 0'143(70'11-55) = 167'838$$

$$\alpha: [70'11,\ 167'838]$$

d: Trennlinie A2 $[68, 158]$ $\qquad d = \dfrac{|68m + 158 + b|}{\sqrt{m^2+1}}$ (3)

d: Trennlinie B2 $[75, 190]$ $\qquad d = \dfrac{|75m + 190 + b|}{\sqrt{m^2+1}}$ (4)

Trennlinie geht durch α: $y_\alpha = m \cdot x_\alpha + b$

$167{'}838 = 70{'}11\, m + b$ (5)

(3) = (4): $68m + 158 + \cancel{b} = 75m + 190 + \cancel{b}$

$\rightarrow m = -4{'}57$

(5): $167{'}838 = 70{'}11 \cdot (-4{'}57) + b \rightarrow$

$b = 167{'}838 + 70{'}11 \cdot 4{'}57 \longrightarrow b = 488{'}2407$

$\boxed{y = -4{'}57\, x + 488{'}2407}$

$\boxed{d = 71{'}71 \;\; (3)}$

$\boxed{\text{Supportvektoren A2, B2}}$