Main exercise in Machine Learning

K-means

iterația:
→ calcul μ → 1)
→ calcul C → 2)

Fie setul de date

	x_1	x_2
A	-1	0
B	1	0
C	0	1
D	3	0
E	3	1

Aplic alg 2-means (k), folosind urm inițializ:

$$\mu_1^0 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

$$\mu_2^0 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Metoda 1 (analitică): merge pt orice caz

- calc dist de la fiecare pct la fiecare centroid

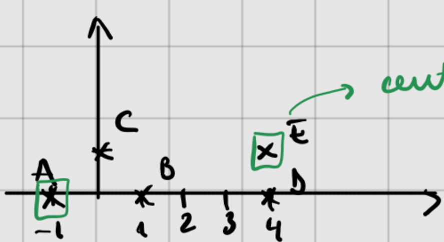
$d(-,-)$	A	B	C	D	E
μ_1^0	0	2	$\sqrt{2}$	4	$\sqrt{17}$
μ_2^0	$\sqrt{17}$	$\sqrt{5}$	3	1	0

mai mică dist față de centroid

$$\| \mu_1^0 - A \| = \left\| \begin{bmatrix} -1 \\ 0 \end{bmatrix} - \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\| = 0$$

$$\| \mu_2^0 - A \| = \left\| \begin{bmatrix} 3 \\ 1 \end{bmatrix} - \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 4 \\ 1 \end{bmatrix} \right\| = \sqrt{16+1} = \sqrt{17}$$

calc rapid vizual:



$\Rightarrow \quad C_1^0 = \{A, B, C\}$

$C_2^0 = \{D, E\}$

iterație $\to$ 1

calc. centroizii

$$\mu_1^1 = \frac{A+B+C}{3} \overset{\text{medie vect}}{=} \frac{\begin{bmatrix}-1\\0\end{bmatrix}+\begin{bmatrix}1\\0\end{bmatrix}+\begin{bmatrix}0\\1\end{bmatrix}}{3} = \frac{\begin{bmatrix}0\\1\end{bmatrix}}{3} = \begin{bmatrix}0\\\frac{1}{3}\end{bmatrix}$$

$$\mu_2^1 = \frac{D+E}{2} = \frac{\begin{bmatrix}3\\0\end{bmatrix}+\begin{bmatrix}3\\1\end{bmatrix}}{2} = \begin{bmatrix}3\\0,5\end{bmatrix}$$

	A	B	C	D	E
μ_1^1	1,05	1.05	0,66	3,01	3,07
μ_2^1	4.03	2,06	3,04	0,5	0,5

$$\| \mu_1^1 - A \| = \left\| \begin{bmatrix}0\\\frac{1}{3}\end{bmatrix} - \begin{bmatrix}-1\\0\end{bmatrix} \right\| = \ldots$$

$\Downarrow$

$C_1^1 = \{A, B, C\}$

$C_2^1 = \{D, E\}$

$\to$ dacă sunt egale la it consecutive ne oprim

$C_1^0 = C_1^1$

$C_2^0 = C_2^1$ $\quad \Big| \to \boxed{\text{STOP!}}$

K means Criteriul J

d.k funcții pe cazul general

set de date

X
-9
-8
-7
-6
-5
5
5
5
6
6
7
7
8
8
8
9
9

Considerăm că aplicăm alg 2-means cu urm inițializare:

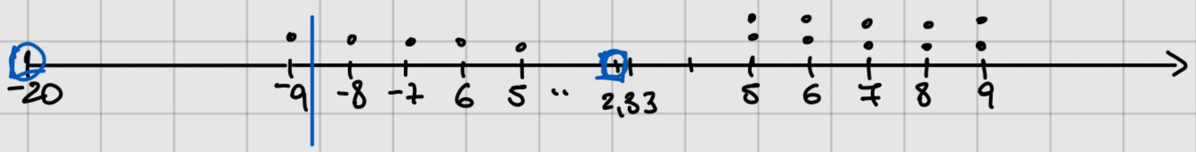
$$\begin{cases} \mu_1^0 = -20 \\ \mu_2^0 = \frac{7}{3} \approx 2,33 \end{cases}$$

Dem în manieră analitică (NU numeric) că:

$$J(C^0, \mu^0) \geq J(C^1, \mu^1)$$

it = 0 : $\mu_1^0 = -20$; $C_1^0 = \{-9\}$

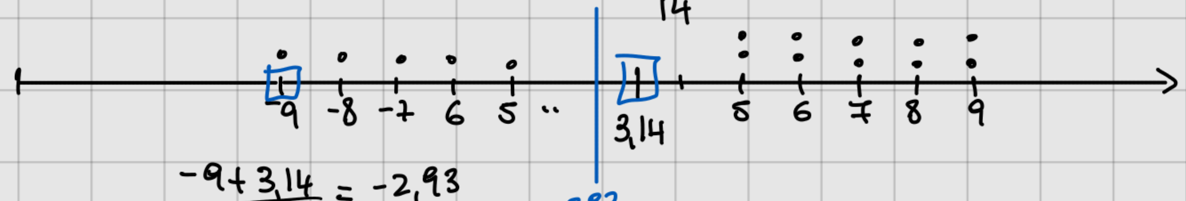
$\mu_2^0 = 7/3$; $C_2^0 = \{-8, \dots -5, 5, \dots 9\}$



$$\frac{-20 + 2{,}33}{2} = -8{,}8$$

it = 1 :

$$\begin{cases} \mu_1^1 = \frac{-9}{1} = -9 \\ \mu_2^1 = \frac{-8-7-6-5+2(5+6+7+8+9)}{14} = \frac{22}{7} \approx 3{,}14 \end{cases}$$



$$\frac{-9 + 3{,}14}{2} = -2{,}93$$

$C_1^1 = \{-9, -8, -7, -6, -5\}$

$C_2^1 = \{5, 5, 6, 6, 7, 7, 8, 8, 9, 9\}$

Vom dem $J(C^0, \mu^0) \geq J(C^1, \mu^1)$ arătând că:

$$J(C^0, \mu^0) \overset{(1)}{\geq} J(C^0, \mu^1) \overset{(2)}{\geq} J(C^1, \mu^1)$$

mai întâi μ
după actualizăm C

Ineg (1): $J(C^0, \mu^0) = \|-9-(-20)\|^2 + \|-8 - \frac{7}{3}\|^2 + \|-7 - \frac{7}{3}\|^2 + \dots + \|9 - \frac{7}{3}\|^2$

$$J(C^0, \mu^1) = \|-9 - (-9)\|^2 + \|-8 - \tfrac{22}{7}\|^2 + \|-7 - \tfrac{22}{7}\|^2 + \dots + \|9 - \tfrac{22}{7}\|^2$$

$$f(x) = \|-9 - x\|^2$$
$$g(y) = \|-8-y\|^2 + \|-7-y\|^2 + \dots + \|9-y\|^2$$

$$J(C^0, \mu^0) = f(-20) + g(\tfrac{7}{3})$$
$$J(C^0, \mu^1) = f(-9) + g(\tfrac{22}{7})$$

$\cup$ $f(x) = (-9-x)^2 = (9+x)^2 = \underbrace{81}_{c} + \underbrace{18x}_{b} + \underbrace{x^2}_{a}$

$a = 1 > 0 \Rightarrow f$ are minim $\Rightarrow x_{min} = \frac{-b}{2a} = \frac{-18}{2 \cdot 1} = -9$

$g(y) = (-8-y)^2 + \dots + (9-y)^2 =$

$= 64 + 16y + y^2 + 49 + 14y + y^2 + \dots + 81 - 18y + y^2 =$

$= 14y^2 + (-88)y + \dots$

$a = 14 > 0 \Rightarrow g$ are min $\Rightarrow y_{min} = \frac{-b}{2a} = \frac{-88}{2 \cdot 14} = \frac{22}{7}$

$J(c^0, \mu^0) = f(-20) + g(\frac{7}{3})$

$J(c^0, \mu^1) = \quad (-9) + g(\frac{22}{7})$

$\Rightarrow J(c^0, \mu^0) > J(c^0, \mu^1)$ (1)

Ineg (2): $J(c^0, \mu^1) \geq J(c^1, \mu^1)$

$J(c^0, \mu^1) = \|-9-(-9)\|^2 + \|-8 - \frac{22}{7}\|^2 + \dots + \|9 - \frac{22}{7}\|^2$

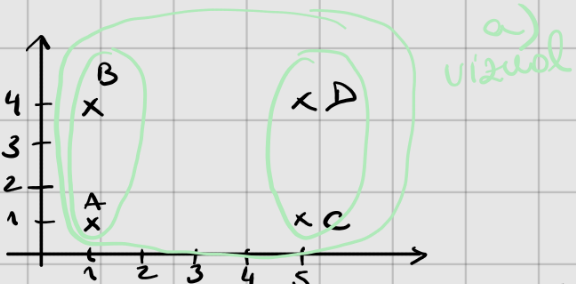
$J(c^1, \mu^1) = \|-9 - (-9)\|^2 + \|\underbrace{-8-(-9)}_{1^2}\|^2 + \dots \|-5-(-9)\|^2 + \|9 - \frac{22}{7}\|^2$

* Comparăm termen cu termen $\Rightarrow J(c^0, \mu^1) > J(c^1, \mu^1)$ (2)

Din (1) (2) $\Rightarrow J(c^0, \mu^0) > J(c^1, \mu^1)$

Clusterizare ierarhică

	x_1	x_2
A	1	1
B	1	4
C	5	1
D	5	4



a) Aplicați alg de clusterizare ierarhică bottom-up folosind distanța euclidiană între vectori și distanța single-linkage ca distanță între clustere

b) -ll- complete linkage -ll-

c) -ll- average linkage -ll-

d) -ll- lui Ward -ll-

a)

	A	B	C	D
A	0			
B	3	0		
C	4	5	0	
D	5	4	3	0

$d_{sl}(B, A) =$

$= \min(d_{sl}(B, A)) =$

$= d(B, A) =$

$= d_2(B, A) =$

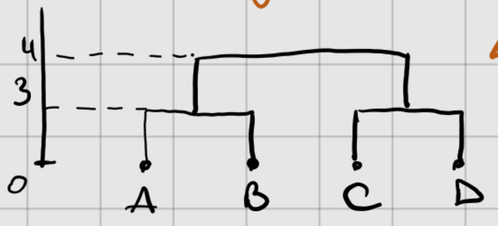
↳ euclidiană

convenție: ordine alfabetică

$= d_2\left(\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\4\end{bmatrix}\right) =$

$= \sqrt{(1-1)^2+(1-4)^2} = \sqrt{9} = 3$

dendrogramă finală:



	A|B	C	D
A|B	0		
C	4	0	
D	4	3	0

$d_{sl}(A|B, C) = \min(d_2(A,C), d_2(B,C)) = 4$

$d_{sl}(A|B, D) = \min(d_2(A,D), d_2(B,D)) = 4$

$d_{sl}(C,D) = d_2(C,D) = 3$

	AB	C|D
AB	0	
C|D	4	0

$d_{sl}(AB, C|D) =$

$\overset{def}{=} \min\begin{pmatrix} d(A,C), & d(A,D) \\ d(B,C), & d(B,D)\end{pmatrix} = 4$

sau

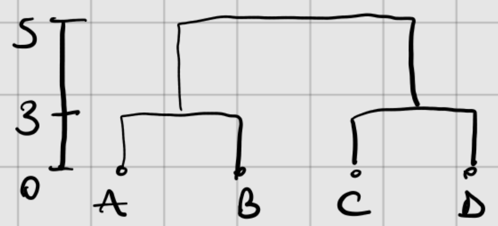
$\overset{prop}{=} \min(d_{sl}(AB,C), d_{sl}(AB,D)) = 4$

b) ...

$d_{cl}(AB, C|D) =$

$\overset{def}{=} \max(d_2(A,C), d_2(A,D), d_2(B,C), d_2(B,D)) = 5$

$\overset{prop}{=} \max(d_{cl}(AB,C), d_{cl}(AB,D)) = 5$



c) ...

$d_{al}(AB, C|D) =$

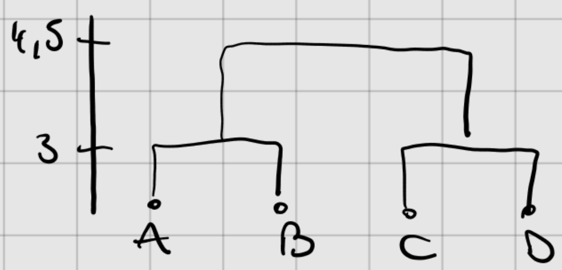
$\overset{def}{=} \dfrac{d_2(A,C)+d(A,D)+d_2(B,C)+d_2(B,D)}{4}$

$= \dfrac{2\cdot 4 + 2\cdot 5}{4} = 4{,}5$

$\overset{prop}{=} \dfrac{\cancel{|AB|}|C|\, d_{al}(AB,C) + |AB||D|\, d_{al}(AB,D)}{\cancel{|AB|}\, \underbrace{|CD|}_{|C|+|D|}}$

$= \dfrac{d_{al}(AB,C) + d_{al}(AB,D)}{2}$

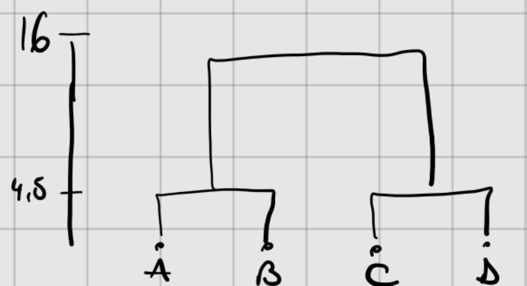
$= \dfrac{4{,}5+4{,}5}{2} = 4{,}5$



d) ...

$d_{ward}(A,B) =$

$= \|A - \mu_{AB}\|^2 +$
$+ \|B - \mu_{AB}\|^2 -$
$- \|A - \mu_A\|^2 -$
$- \|B - \mu_B\|^2$



$$AB = \frac{A+B}{2} = \frac{\begin{bmatrix}1\\1\end{bmatrix}+\begin{bmatrix}1\\4\end{bmatrix}}{2} = \frac{\begin{bmatrix}2\\5\end{bmatrix}}{2} = \begin{bmatrix}1\\2,5\end{bmatrix}$$

$$\mu_A = \frac{A}{1} = A \qquad \mu_B = \frac{B}{1} = B$$

$$= \left\| \begin{bmatrix}1\\1\end{bmatrix} - \begin{bmatrix}1\\2,5\end{bmatrix} \right\|^2 + \left\| \begin{bmatrix}1\\4\end{bmatrix} - \begin{bmatrix}1\\2,5\end{bmatrix} \right\|^2$$

$$= \left\| \begin{bmatrix}0\\-1,5\end{bmatrix} \right\|^2 + \left\| \begin{bmatrix}0\\1,5\end{bmatrix} \right\|^2 = (-1,5)^2 + (1,5)^2 = 2 \cdot 1,5 \approx 4,5$$

SAU

$$= \frac{|A|\,|B|}{\underbrace{|AB|}_{|A|+|B|}} \cdot \| \underset{A}{\mu_A} - \underset{B}{\mu_B} \|^2 =$$

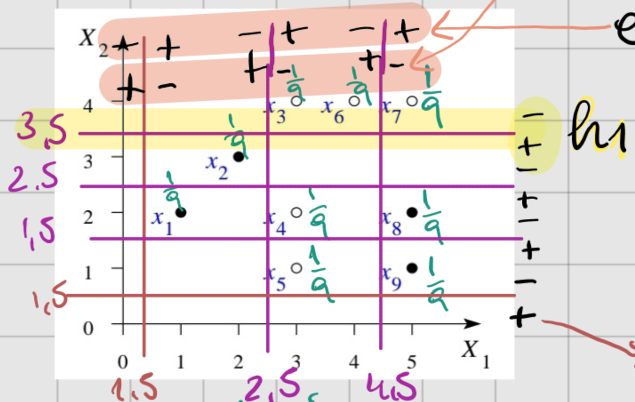
$$= \frac{1 \cdot 1}{1+1} \cdot 9 = \frac{9}{2} = 4,5$$

AdaBoost

x_i	X_1	X_2	y_i
x_1	1	2	+1
x_2	2	3	+1
x_3	3	4	−1
x_4	3	2	−1
x_5	3	1	−1
x_6	4	4	−1
x_7	5	4	−1
x_8	5	2	+1
x_9	5	1	+1

$t = 1$

s	0,5	2,5	4,5
$err_{D_1}(X_1 < s)$	$\frac{4}{9}$	$\frac{2}{9}$	$\frac{4+2}{9} = \frac{6}{9}$
$err_{D_1}(X_1 \geq s)$	$\frac{5}{9}$	$\frac{7}{9}$	$\frac{3}{9}$



−2 compose pt fiecare split (+, −)

$1 - \frac{4}{9} = \frac{5}{9}$

când e mai mic punem +

- ponderi
- trebuie să luăm un sep. exterior mereu (stg sau dr)

s	0,5	1,5	2,5	3,5
$er_{D_1}(X_2 < 1)$	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{3}{9}$	$\frac{2}{9}$
$er_{D_1}(X_2 \geq 1)$	$\frac{5}{9}$	$\frac{5}{9}$	$\frac{6}{9}$	$\frac{7}{9}$

→ minimul

$\varepsilon_1 = \frac{2}{9}$ (luăm minimul din tabel)

$h_1 = (x_2 < 3{,}5) = \text{sign}(3{,}5 - x_2)$

ipoteza pe care o alege

$\gamma_1 = \frac{1}{2} - \varepsilon_1 = \frac{5}{18}$

$\alpha_1 = \frac{1}{2} \ln \frac{1-\varepsilon_1}{\varepsilon_1} = \frac{1}{2} \ln \frac{1-\frac{2}{9}}{\frac{2}{9}} = \frac{1}{2} \ln \frac{7}{2} = \ln \sqrt{\frac{7}{2}} =$
$= 0{,}626$

ipoteza la finalul rulării

$\hookrightarrow H_1(x) = \text{sign}(\alpha_1 h_1(x)) =$
$= \text{sign}(0{,}626 \cdot \text{sign}(3{,}5 - x \cdot x_2))$

Calcul D_2:

metoda 1:

$e^{\sqrt{\frac{7}{2}}^{-1}} = \frac{2}{7}$

etichete reale

etichete aleasă de AdaB

$D_2(i) = \frac{1}{z_1} \underbrace{D_1(i)}_{\frac{1}{9}} \left(e^{-\alpha_1}\right)^{y_i h_1(x_i)}$

-1	-1	⇒ 1
1	1	⇒ 1
-1	1	⇒ -1
1	-1	⇒ -1

⇒ $\begin{cases} 1, & h_1 \text{ clasifică corect } x_i \\ -1, & \text{— greșit } x_i \end{cases}$

$= \begin{cases} \frac{1}{z_1} \cdot \frac{1}{9} \cdot \left(\sqrt{\frac{2}{7}}\right)^{-1}, & i \in \{4,5\} \rightarrow \text{greșit} \\ \frac{1}{z_1} \cdot \frac{1}{9} \cdot \left(\sqrt{\frac{2}{7}}\right)^{1}, & i \in \{1,2,3,6,7,8,9\} \rightarrow \text{corect} \end{cases}$

nu știm z_1; ce înseamnă fct de normaliz.?

PROP $\sum_{i=1}^{9} D_2(i) = 1 \Rightarrow 7 \cdot \frac{1}{z_1} \cdot \frac{1}{9} \cdot \sqrt{\frac{2}{7}} + \frac{1}{z_1} \cdot \frac{1}{9} \cdot \sqrt{\frac{7}{2}} \cdot 2 = 1 \,/ \cdot z_1$

$\Rightarrow \frac{7}{9}\sqrt{\frac{2}{7}} + \frac{2}{9}\sqrt{\frac{7}{2}} = z_1$

$\Rightarrow z_1 = \frac{1}{9}\left(\sqrt{14} + \sqrt{14}\right) = \frac{2\sqrt{14}}{9}$

Avem z_1 și putem înlocui

$= \begin{cases} \frac{1}{4}, & i \in \{4,5\} \\ \frac{1}{14}, & i \notin \{4,5\} \end{cases}$

metodo 2:

corect clasific. : $i \in \{1,2,3,6,7,8,9\}$: $\overbrace{\frac{1}{9}\cdot a}^{D_2(1)} + \overbrace{\frac{1}{9}\cdot a}^{D_2(2)} + \overbrace{\frac{1}{9}\cdot a}^{D_2(3)} + \overbrace{\frac{1}{9}\cdot a}^{D_2(6)} + \overbrace{\frac{1}{9}\cdot a}^{D_2(7)} + \overbrace{\frac{1}{9}\cdot a}^{D_2(8)} + \overbrace{\frac{1}{9}\cdot a}^{D_2(9)} = \frac{1}{2}$

$\frac{1}{14}$

incorect clasif. : $i \in \{4,5\}$: $\frac{1}{9} b + \frac{1}{9} b = \frac{1}{2}$

$\frac{9}{4}$

$\Rightarrow \frac{7}{9} a = \frac{1}{2} \Rightarrow a = \frac{9}{14}$

$\Rightarrow \frac{2}{9} b = \frac{1}{2} \Rightarrow b = \frac{9}{4}$

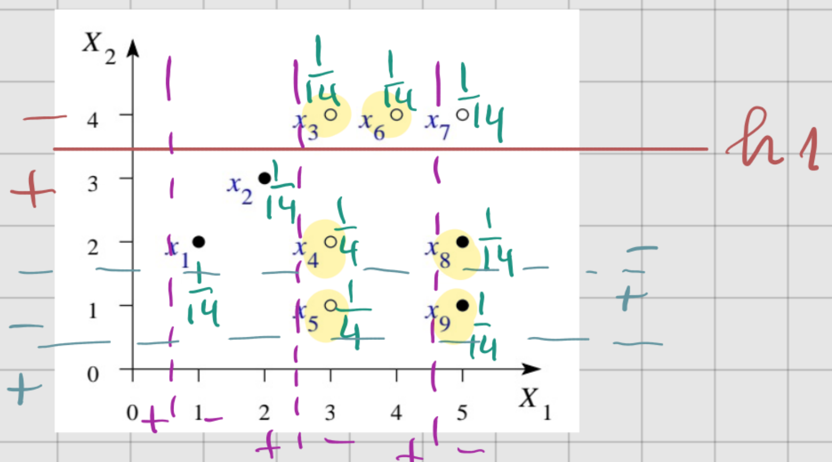
metodo 3

$$D_2(i) = \begin{cases} \dfrac{D_1(i)}{2\cdot\frac{2}{9}}, & i \in \{4,5\} \\[2ex] \dfrac{D_1(i)}{2\cdot\frac{7}{9}}, & i \notin \{4,5\} \end{cases}$$

$\dfrac{\frac{1}{9}}{2\cdot\frac{2}{9}} = \frac{1}{4}$

$\dfrac{\frac{1}{9}}{2\cdot\frac{7}{9}} = \frac{1}{14}$

$$D_2 : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \frac{1}{14} & \frac{1}{14} & \frac{1}{14} & \frac{1}{4} & \frac{1}{4} & \frac{1}{14} & \frac{1}{14} & \frac{1}{14} & \frac{1}{14} \end{pmatrix}$$



$t = 2$

s	0,5	2,5	4,5
$er_{D_2}(x_1 < s)$	$\frac{4}{14}$	$\frac{2}{14}$	$\frac{4}{14} + \frac{2}{4} = \frac{11}{14}$
$er_{D_2}(x_1 \geq s)$	$\frac{10}{14}$	$\frac{12}{14}$	$\frac{3}{14}$

s	0,5	1,5	2,5	3,5
$er_{D_2}(x_2 < s)$	$\frac{4}{14}$	$\frac{3}{14} + \frac{1}{4} = \frac{13}{28}$	$\frac{2}{4} + \frac{1}{14} = \frac{4}{7}$	$\frac{2}{4} = \frac{1}{2}$
$er_{D_2}(x_2 \geq s)$	$\frac{10}{14}$	$\frac{15}{28}$	$\frac{3}{7}$	$\frac{1}{2}$

știm sigur că urm iter. vor fi $\frac{1}{2}$

$$\varepsilon_2 = \frac{2}{14} = \frac{1}{7}$$

$$h_2 = (x_1 < 2,5) = \\ = \text{sign}(2,5 - x_1)$$

$$\gamma_2 = \frac{1}{2} - \varepsilon_2 = \frac{5}{14}$$

$$\alpha_2 = \frac{1}{2} \ln \frac{1-\varepsilon_2}{\varepsilon_2} = \ln\sqrt{6}$$

vrem să știm + sau −

$$H_2(x) = \text{sign}\big(\alpha_1 h_1(x) + \alpha_2 h_2(x)\big)$$

$$= \text{sign}\big(0,626 \cdot \text{sign}(3,5 - x \cdot X_2) + \\ + 0,896 \cdot \text{sign}(2,5 - x \cdot X_1)\big)$$

ipoteza finală dacă alg s-ar fi oprit la această iterație

(pt o nouă instanță înlocuim și clasific la final cu + sau −)

Calcul D3:

metoda 3:

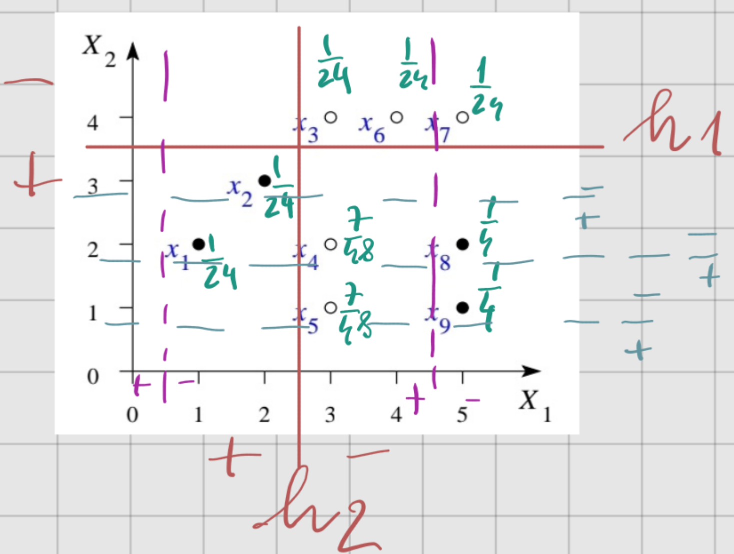
$$D_3(i) = \begin{cases} \dfrac{D_1(i)}{2 \cdot \frac{1}{7}}, & i \in \{8,9\} \\ \dfrac{D_1(i)}{2 \cdot \frac{6}{7}}, & i \notin \{8,9\} \end{cases}$$

ε_2 ← ε_M greșit

ε_2 ← ε_C corecte

$$D_3: \begin{array}{ccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array}$$
$$\begin{pmatrix} \frac{\frac{1}{14}}{2\cdot\frac{6}{7}} & \frac{\frac{1}{14}}{2\cdot\frac{6}{7}} & \frac{\frac{1}{14}}{2\cdot\frac{6}{7}} & \frac{\frac{1}{4}}{2\cdot\frac{6}{7}} & \frac{\frac{1}{4}}{2\cdot\frac{6}{7}} & \frac{\frac{1}{14}}{2\cdot\frac{6}{7}} & \frac{\frac{1}{14}}{2\cdot\frac{6}{7}} & \frac{\frac{1}{14}}{2\cdot\frac{1}{7}} & \frac{\frac{1}{14}}{2\cdot\frac{1}{7}} \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{1}{24} & \frac{1}{24} & \frac{1}{24} & \frac{7}{48} & \frac{7}{48} & \frac{1}{24} & \frac{1}{24} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$



s	0,5	2,5	4,5
$er_{D_3}(x_1<s)$	$\frac{2}{24}+\frac{2}{4}=\frac{8}{24}=\frac{7}{12}$	$\frac{1}{2}$	$\frac{2}{24}+\frac{14}{48}+\frac{2}{4}=\frac{21}{24}$
$er_{D_3}(x_1\geq s)$	$\frac{5}{12}$	$\frac{1}{2}$	$\frac{3}{24}=\frac{1}{8}$ minim

s	0,5	1,5	2,5	3,5
$er_{D_3}(x_2<s)$	$\frac{2}{24}+\frac{2}{4}=\frac{7}{12}$	$\frac{7}{48}+\frac{1}{4}+\frac{2}{24}=\frac{23}{48}$	$\frac{14}{48}+\frac{1}{24}=\frac{1}{3}$	$\frac{14}{48}=\frac{7}{24}$
$er_{D_3}(x_2\geq s)$	$\frac{5}{12}$	$\frac{25}{48}$	$\frac{2}{3}$	$\frac{17}{24}$

$\varepsilon_8 = \frac{1}{8}$

$h_3 = (x_1 \geq 4{,}5) = \text{sign}(x_1 - 4{,}5)$

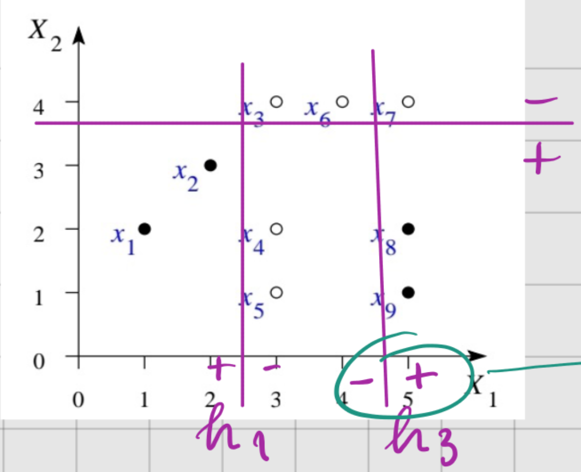
$\gamma_3 = \frac{1}{3} - \varepsilon_8 = \frac{3}{8}$

$\alpha_3 = \ln\sqrt{7} = 0{,}973$

$H_3(x) = \text{sign}(\alpha_1 h_1(x) + \alpha_2 h_2(x) + \alpha_3 h_3(x))$
$= \text{sign}\,(0{,}626 \cdot \text{sign}(3{,}5 - x \cdot x_2) +$
$+ 0{,}896 \cdot \text{sign}(2{,}5 - x \cdot x_1) +$
$+ 0{,}973 \cdot \text{sign}(x \cdot x_1 - 4{,}5))$

t	α_t	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	
1	0.626	+1	+1	−1	+1	+1	−1	−1	+1	+1	$\to h_1$
2	0.896	+1	+1	−1	−1	−1	−1	−1	−1	−1	$\to h_2$
3	0.973	−1	−1	−1	−1	−1	−1	+1	+1	+1	$\to h_3$
	$H_T(x_i)$	+1	+1	−1	−1	−1	−1	−1	+1	+1	$\to H_3$

" sign$(0 \cdot 0{,}626 \cdot (+1) + 0 \cdot 0{,}896\,(+1) + 0{,}973 \cdot (-1))$
$= \quad > 0 \quad = +1$



h_1

$\to$ pt că am ales $er_{D_3}(x_1 \geq s)$

Eroarea la antrenare

x_i	X_1	X_2	y_i
x_1	1	2	+1
x_2	2	3	+1
x_3	3	4	−1
x_4	3	2	−1
x_5	3	1	−1
x_6	4	4	−1
x_7	5	4	−1
x_8	5	2	+1
x_9	5	1	+1

→ compar cu H_3 $\Rightarrow$ er antrenare$_{AB} = 0$

b) $(x_1 = 1, x_2 = 5)$ ← Clasificăm instanța

$H_3(x_1 = 1, x_2 = 5) =$

$= \text{sign}\big(0{,}6\,(-1) + 0{,}8 \cdot (+1) + 0{,}9 \cdot (-1)\big) = <0 = -1$

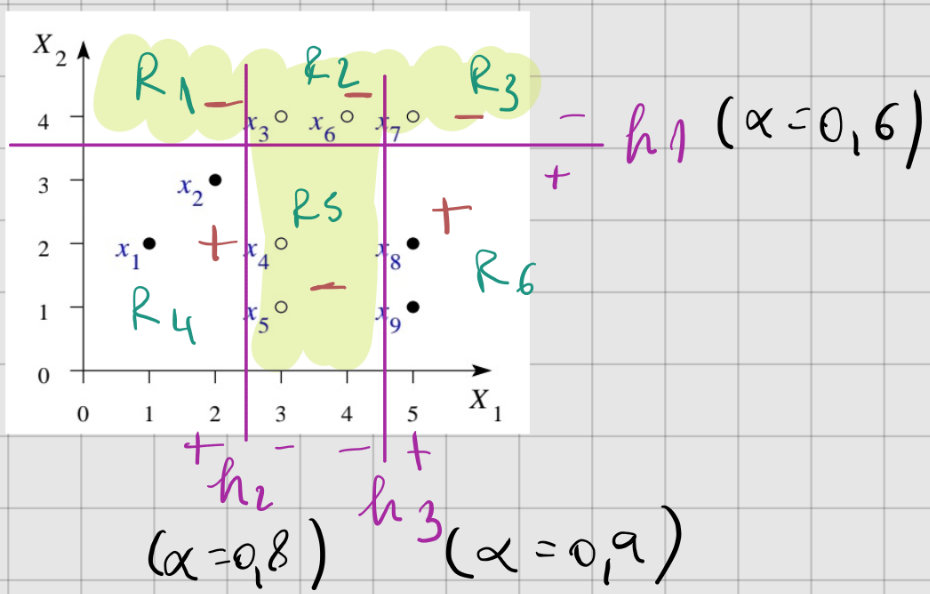
$H_3(x) = \text{sign}(\alpha_1 h_1(x) + \alpha_2 h_2(x) + \alpha_3 h_3(x))$

$= \text{sign}\big(0{,}626 \cdot \text{sign}(3{,}5 - x \cdot x_2) + \quad \xrightarrow{5} \quad 3{,}5 - 5 < 0 = -1$

$+ 0{,}896 \cdot \text{sign}(2{,}5 - x \cdot x_1) + \quad \xrightarrow{1} \quad 2{,}5 - 1 > 0 = +1$

$+ 0{,}973 \cdot \text{sign}(\underbrace{x \cdot x_1 - 4{,}5}_{1})\big) \quad \rightarrow \quad 1 - 4{,}5 < 0 = -1$

c) granițe de decizie



t	α	R1	R2	R3	R4	R5	R6
1	0,6	−	−	−	+	+	+
2	0,8	+	−	−	+	−	−
3	0,9	−	−	+	−	−	+
	$H_3(R_i)$	−	−	−	+	−	+

↳ calculăm

K-NN

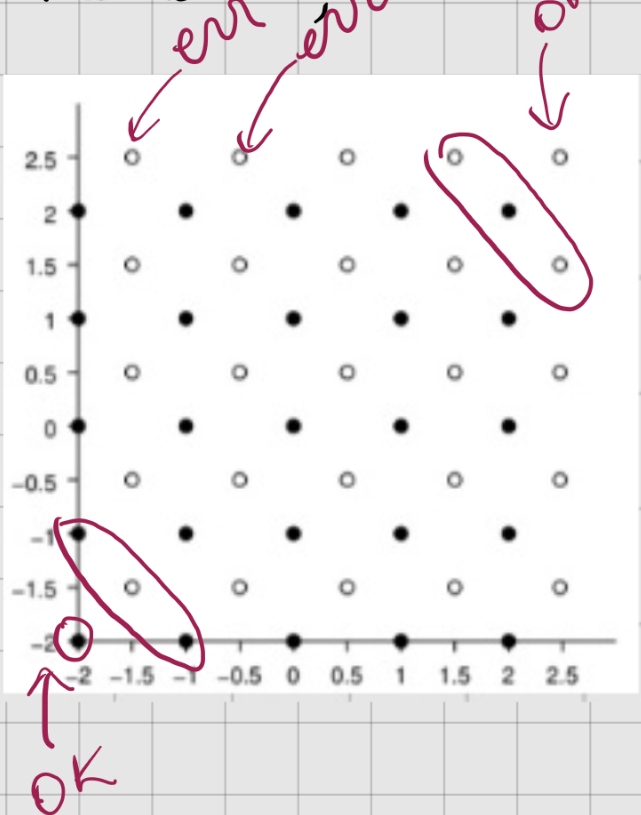
Data	Eticheta	Vecinătate	CVLOO	Eroare
(-2,-2)	1	(-1,5;-1,5)	0	Da
(-1,-2)	1	(-1,5;-1,5)/ (-0,5;-1,5)	0	Da
....				
(-1,5;-1,5)	0	(-2,-2)/(-1,-2)/ (-2,-1)/(-1,-1)	1	Da
...				

err CVLOO = 1

=> 100% eroare la cross-validare
(1-NN este foarte sensibil la mixarea datelor)

Eroarea la antrenare pt 1-NN este 0, întrucât datele de antrenament nu conțin inconsistențe

c)



Data	Eticheta	Vecinătate	CVLOO	Eroare
(-2,-2)	1	(-1,5;-1,5)=0 (-2,-1)/(-1,-2) =1 =1	majoritar 1	Nu
...				
(-1,5,-1,5)	0	(-2,-2)=1/(-1,-2)=1 (-2,-1)=1/(-1,-1)=1 (-0,5,-1,5)/=0 (-1,5,-0,5)=0	1	Da

=> eroare CVLOO $= \frac{48}{50}$

Regresia Logistică

x_0	x_1	x_2	y	$y^{(0)}$	$y^{(1)}$
1	0	0	1	0,5	0,5
1	0	1	0	0,5	0,51
1	1	0	0	0,5	0,54
1	2	2	1	0,5	0,54

1) metoda gradientului ascendent $w^{(0)} = \begin{bmatrix}0\\0\\0\end{bmatrix}$ $\alpha = 0,1$

2) met. lui Newton $w^{(0)} = \begin{bmatrix}0\\0\\0\end{bmatrix}$

$\sigma = \frac{1}{1+e^{-x}}$ $\begin{pmatrix} 0 & 1 \\ 1-\sigma(wx) & \sigma(wx) \end{pmatrix}$

$Y|X \sim \text{Bernoulli}\left(\sigma\left(w \cdot x^{(i)}\right)\right)$

$w \cdot x \rightarrow \begin{bmatrix}1\\2\end{bmatrix}\begin{bmatrix}3\\4\end{bmatrix} = 1\cdot 3 + 2 \cdot 4 \rightarrow$ produs scalar

a) $$w^{(t)} \leftarrow w^{(t-1)} + \underbrace{\alpha \sum_{i=1}^{n} \overbrace{\left(y^{(i)} - \sigma\left(w^{(t-1)} \cdot x^{(i)}\right)\right)}^{y} x^{(i)}}_{\frac{\partial l_D}{\partial w}\left(w^{(t-1)}\right)}$$

$$w^{(1)} = w^{(0)} + \alpha \cdot \frac{\partial l_D}{\partial w}\left(w^{(0)}\right)$$

$$\frac{\partial l_D}{\partial w}\left(w^{(0)}\right) = (1-0,5)\begin{bmatrix}1\\0\\0\end{bmatrix} + (0-0,5)\begin{bmatrix}1\\0\\1\end{bmatrix} + (0-0,5)\begin{bmatrix}1\\1\\0\end{bmatrix} + (1-0,5)\begin{bmatrix}1\\2\\2\end{bmatrix}$$
$$= \begin{bmatrix}0,5\\0\\0\end{bmatrix} + \begin{bmatrix}-0,5\\0\\-0,5\end{bmatrix} + \begin{bmatrix}-0,5\\-0,5\\0\end{bmatrix} + \begin{bmatrix}0,5\\1\\1\end{bmatrix} = \begin{bmatrix}0\\0,5\\0,5\end{bmatrix}$$

↳ este GRADIENT

$w^{(0)} \cdot x^{(1)} = \begin{bmatrix}0\\0\\0\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix} = 0$

$\sigma\left(w^{(0)} \cdot x^{(1)}\right) = \sigma(0) = \frac{1}{1+e^{-0}} = \frac{1}{1+1} = 0,5$

$w^{(0)} \cdot x^{(2)} = \begin{bmatrix}0\\0\\0\end{bmatrix}\begin{bmatrix}1\\0\\1\end{bmatrix} = 0$

...

$$w^{(1)} = \begin{bmatrix}0\\0\\0\end{bmatrix} + 0,1\begin{bmatrix}0\\0,5\\0,5\end{bmatrix} = \begin{bmatrix}0\\0,05\\0,05\end{bmatrix}$$

$$l_D(w) = \sum_{i=1}^{n}\left(\left(1-y^{(i)}\right)\ln\left(1-\sigma\left(w \cdot x^{(i)}\right)\right) + y^{(i)}\ln\left(\sigma\left(w \cdot x^{(i)}\right)\right)\right)$$

$y^{(i)} = 0$

$y^{(i)} = 1$

$y: \begin{matrix}1\\0\\0\\1\end{matrix}$

$$l_D\left(w^{(0)}\right) = \ln\left(\sigma\left(w^{(0)} \cdot x^{(0)}\right)\right) + \ln\left(1-\sigma\left(w^{(0)} \cdot x^{(1)}\right)\right) + \ln\left(1-\sigma\left(w^{(0)} \cdot x^{(2)}\right)\right) + \ln\left(\sigma\left(w^{(0)} \cdot x^{(3)}\right)\right)$$

$$= \ln(0,5) + \ln(0,5) + \ln(0,5) + \ln(0,5) = 4\ln(0,5) = -2,7725$$

$$\ell_D(w^{(1)}) = \ln(\sigma(w^{(1)}\cdot x^{(0)})) + \ln(1-\sigma(w^{(1)}\cdot x^{(1)})) + \ln(1-\sigma(w^{(1)}\cdot x^{(2)})) + \ln(\sigma(w^{(1)}\cdot x^{(3)}))$$

$$w^{(1)}\cdot x^{(1)} = \begin{bmatrix}0\\0,05\\0,05\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix} = 0 \Rightarrow \sigma(w^{(1)}x^{(1)}) = \sigma(0) = 0,5$$

$$w^{(1)}\cdot x^{(2)} = \begin{bmatrix}0\\0,05\\0,05\end{bmatrix}\begin{bmatrix}1\\0\\1\end{bmatrix} = 0,05 \Rightarrow \sigma(0,05) = 0,5124$$

$$w^{(1)}\cdot x^{(3)} = \begin{bmatrix}0\\0,05\\0,05\end{bmatrix}\begin{bmatrix}1\\1\\0\end{bmatrix} = 0,05 \Rightarrow \sigma(0,05) = 0,5124$$

$$w^{(1)}\cdot x^{(4)} = \begin{bmatrix}0\\0,05\\0,05\end{bmatrix}\begin{bmatrix}1\\2\\2\end{bmatrix} = 0,2 \Rightarrow \sigma(0,2) = 0,5498$$

$$\ell_D(w^{(1)}) = \ln(0,5) + \ln(1-0,5124)\cdot 2 + \ln(0,5498) = -2,7278$$

$$\boxed{\underset{-2,77}{\ell_D(w^{(0)})} < \underset{-2,72}{\ell_D(w^{(1)})}}$$

$w^{(2)} \leftarrow - \dots -$

până ce ajungem la convergență

după alg $\longrightarrow$ it $= 244$

$$w^{(244)} \approx w^{(243)} \Rightarrow \text{STOP}$$

$$\Rightarrow w_{MLE} = w^{(244)} = \begin{bmatrix}-0,7043\\0,4880\\0,4880\end{bmatrix}$$

(2) Newton

$$\frac{\partial \ell_D}{\partial w}(w^{(0)}) = \begin{bmatrix}0\\0,5\\0,5\end{bmatrix}$$

Mat. Hessiana:

$$H(w^{(0)}) = -\left(0,5(1-0,5)\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix} + 0,5(1-0,5)\begin{bmatrix}1&0&1\\0&0&0\\1&0&1\end{bmatrix} + 0,5(1-0,5)\begin{bmatrix}1&1&0\\1&1&0\\0&0&0\end{bmatrix} + 0,5(1-0,5)\begin{bmatrix}1&2&2\\2&4&4\\2&4&4\end{bmatrix}\right)$$

$$w^{(0)}\cdot x^{(0)} = \begin{bmatrix}0\\0\\0\end{bmatrix}x^{(0)} = 0 \Rightarrow \sigma(\dots) = 0,5$$

x_0	x_1	x_2	y	$y^{(0)}$
1	0	0	1	0,5
1	0	1	0	0,5
1	1	0	0	0,5
1	2	2	1	0,5

$$\begin{bmatrix}1\\0\\0\end{bmatrix}\begin{bmatrix}1&0&0\end{bmatrix} = \begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}$$

$$H(w^{(0)}) = -\left(\begin{bmatrix} 0{,}25 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0{,}25 & 0 & 0{,}25 \\ 0 & 0 & 0 \\ 0{,}25 & 0 & 0{,}25 \end{bmatrix} + \begin{bmatrix} 0{,}25 & 0{,}25 & 0 \\ 0{,}25 & 0{,}25 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0{,}25 & 0{,}5 & 0{,}5 \\ 0{,}5 & 1 & 1 \\ 0{,}5 & 1 & 1 \end{bmatrix} \right)$$

$$= \begin{bmatrix} -1 & -0{,}75 & -0{,}75 \\ -0{,}75 & -1{,}25 & -1 \\ -0{,}75 & -1 & -1{,}25 \end{bmatrix}$$

$$\left(H(w^{(0)})\right)^{-1} = \frac{1}{\det(H(w^{(0)}))} \cdot H(w^{(0)})^*$$

$$H(w^{(0)})^T = \begin{bmatrix} -1 & -0{,}75 & -0{,}75 \\ -0{,}75 & -1{,}25 & -1 \\ -0{,}75 & -1 & -1{,}25 \end{bmatrix}$$

↳ Trebuie să fie simetric / egal TOT TIMPUL

$$H(w^{(0)})^* = \begin{pmatrix} (-1)^{1+1}\begin{vmatrix} -1{,}25 & -1 \\ -1 & -1{,}25 \end{vmatrix} & (-1)^{1+2}\begin{vmatrix} -0{,}75 & -1 \\ -0{,}75 & -1{,}25 \end{vmatrix} & (-1)^{1+3}\begin{vmatrix} -0{,}75 & -1{,}2 \\ -0{,}75 & -1 \end{vmatrix} \\ (-1)^{2+1}\begin{vmatrix} -0{,}75 & -0{,}75 \\ -1 & -1{,}25 \end{vmatrix} & (-1)^{2+2}\begin{vmatrix} -1 & -0{,}75 \\ -0{,}75 & -1{,}25 \end{vmatrix} & (-1)^{2+3}\begin{vmatrix} -1 & -0{,}75 \\ -0{,}75 & -1 \end{vmatrix} \\ (-1)^{3+1}\begin{vmatrix} -0{,}75 & -0{,}75 \\ -1{,}25 & -1 \end{vmatrix} & (-1)^{3+2}\begin{vmatrix} -1 & -0{,}75 \\ -0{,}75 & -1 \end{vmatrix} & (-1)^{3+3}\begin{vmatrix} -1 & -0{,}75 \\ -0{,}75 & -1{,}25 \end{vmatrix} \end{pmatrix}$$

$$= \begin{bmatrix} 0{,}5625 & -0{,}1875 & -0{,}1875 \\ -0{,}1875 & 0{,}6875 & -0{,}4375 \\ -0{,}1875 & -0{,}4325 & 0{,}6875 \end{bmatrix}$$

$$\det(H(w^{(0)})) = (-1)(-1{,}25)(-1{,}25) + (-0{,}75)(-1)(-0{,}75) + (-0{,}75)(-1)(-0{,}75) - (-0{,}75)(-0{,}75)(-1{,}25)) - (-1)(-1)(-1) - (-0{,}75)(-0{,}75)(-1{,}25) =$$

$$= -0{,}28125$$

$$H(w^{(0)})^{-1} = \frac{1}{\det} H^* = \begin{bmatrix} -2 & 0{,}6666 & 0{,}6666 \\ 0{,}6666 & -2{,}4444 & 1{,}5555 \\ 0{,}6666 & 1{,}5555 & -2{,}4444 \end{bmatrix}$$

$$H(w^{(0)})^{-1} \cdot \frac{\partial l_D}{\partial w}(w^{(0)}) = \begin{bmatrix} 0{,}6666 \\ -0{,}4444 \\ -0{,}4444 \end{bmatrix}$$

↳ vectorul pt a calc $w^{(1)}$

$$\underline{w^{(1)}} \leftarrow \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0,6666 \\ -0,4444 \\ -0,4444 \end{bmatrix} = \begin{bmatrix} -0,6666 \\ 0,4444 \\ 0,4444 \end{bmatrix}$$

$l_D(w^{(0)}) = -2,7725$ (era calc)

$$l_D(w^{(1)}) = \ln\sigma\left(w^{(1)} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) + \ln\left(1-\sigma\left(w^{(1)}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right)\right)$$
$$+ \ln\left(1-\sigma\left(w^{(1)}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right)\right) + \ln\left(\sigma\left(w^{(1)}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}\right)\right)$$

$0,66 - 0,44$

$$= \ln\sigma(-0,6666) + \ln(1-\sigma(-0,2222)) +$$
$$+ \ln(1-\sigma(-0,2222)) + \ln(\sigma(1,1111))$$

$$= -2,5420$$

$$\boxed{\underbrace{l_D(w^{(0)})}_{-2,77} < \underbrace{l(w^{(1)})}_{-2,54}}$$

$w^{(2)} \leftarrow$

⋮

până la convergență

→ Avantaj: nu avem α learning rate

dar avem deriv a 2-a ⇒ mai multe calcule

– dar nr de it e mai puțin

... it = 4

$w^{(4)} \approx w^{(3)}$ ⇒ STOP

$$\downarrow$$

$$w_{MLE} = \begin{bmatrix} -0,7044 \\ 0,4880 \\ 0,4880 \end{bmatrix}$$

↳ e la fel ca la ascendent

(altfel era GREȘIT)

↑ Antrenare

↓ Predicție (clasificare)

$x_0 = 1$

$(x_1 = 1, x_2 = 2)$

$(x_1 = 0,5, x_2 = 0,5)$

$x_0 = 1$

$$w_{LogR} = \begin{bmatrix} -0,7044 \\ 0,4880 \\ 0,4880 \end{bmatrix}$$

$$\hat{y}_{logR} = \underset{y \in \{0,1\}}{\operatorname{argmax}} \; \underbrace{P(Y = y \mid X = x)}_{P_y} =$$

$$P_0 = 1 - \sigma(w_{logR} \cdot x) = 0{,}1113$$

$$P_1 = \sigma(w_{logR} \; x) = 0{,}8887$$

$$\sigma(w_{logR} x) = \sigma\left(\begin{bmatrix} -0{,}7044 \\ 0{,}4880 \\ 0{,}4880 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}\right) =$$

$$= \sigma(2{,}0783) = 0{,}8887$$

$$\Rightarrow \hat{y}_{logR} = \underline{1}$$

$$\hat{y}_{logR} = \underset{y \in \{0,1\}}{\operatorname{argmax}} \; P(\underbrace{Y = y \mid X_1 = 0{,}5,\; X_2 = 0{,}5}_{P_y})$$

$$P_0 = 1 - \sigma\left(w_{logR} \begin{bmatrix} 1 \\ 0{,}5 \\ 0{,}5 \end{bmatrix}\right) = 0{,}5539$$

$$P_1 = \sigma\left(w_{logR} \begin{bmatrix} 1 \\ 0{,}5 \\ 0{,}5 \end{bmatrix}\right) = \sigma\left(\begin{bmatrix} -0{,}7044 \\ 0{,}4880 \\ 0{,}4880 \end{bmatrix} \begin{bmatrix} 1 \\ 0{,}5 \\ 0{,}5 \end{bmatrix}\right)$$

$$= \sigma(-0{,}2164) = 0{,}4461$$

$$\Rightarrow \hat{y}_{logR} = \underline{0}$$

Metodo 2

$$P_1 > P_0 \iff \sigma(wx) > 1 - \sigma(wx)$$

$$\iff \frac{1}{1+e^{-wx}} > 1 - \frac{1}{1+e^{-wx}} \quad \Big| \cdot \underbrace{(1+e^{-wx})}_{>0}$$

$$\iff 1 + e^{-wx} > e^{-wx} \quad | \ln$$

$$\iff \ln 1 > -wx$$

$$\iff 0 > -wx \quad | (-1) \iff \boxed{wx > 0}$$

$$\boxed{\begin{aligned} P_1 > P_0 &\iff w_{logR} \cdot x > 0 \\ P_1 = P_0 &\iff w_{logR}\, x = 0 \\ P_1 < P_0 &\iff w_{logR}\, x < 0 \end{aligned}}$$

$$y_{logR} = \begin{cases} 1, & w_{logR} \cdot x \geq 0 \\ 0, & w_{logR}\, x < 0 \\ \text{conventie}, & w_{logR}\, x = 0 \end{cases}$$

I) $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \to \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \to w_{logR} \cdot \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = 2{,}0783 > 0$

$\Rightarrow y_{logR} = 1$

II) $\begin{bmatrix} 0{,}5 \\ 0{,}5 \end{bmatrix} \to \begin{bmatrix} 1 \\ 0{,}5 \\ 0{,}5 \end{bmatrix} \to w_{logR} \begin{bmatrix} 1 \\ 0{,}5 \\ 0{,}5 \end{bmatrix} = -0{,}2164 < 0$

$\Rightarrow y_{logR} = 0$

---

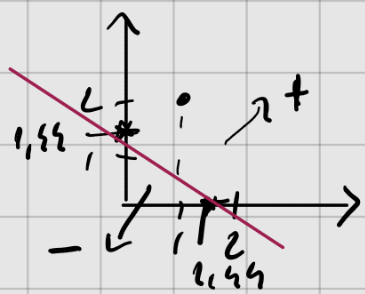
c) Trasăm granița de decizie

$wX = 0 \iff \begin{bmatrix} -0{,}7044 \\ 0{,}4880 \\ 0{,}4880 \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix} = 0$

$\iff -0{,}7044 + 0{,}4880\, x_1 + 0{,}4880\, x_2 = 0$

$x_1 = 0 \Rightarrow x_2 = \dfrac{0{,}7044}{0{,}4880} = 1{,}44$

$x_2 = 0 \Rightarrow x_1 = 1{,}44$



luăm pct pt a vedea semnul

dar din b) $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \to +$