- modelul prumiste population = 10000 oameni la iesire primim 1900 în setul de date: pentru population = 10000 profit = 1200 00

y

$$\left(1900 - (2900\ 00)\right)^2$$

- modelul dă $1900 \to$ $10000 - 1000\ 0 = \boxed{8000}$
  realitate: 10000

- modelul dă $10000 \to$ $10000 - 1900 = \boxed{8000}$
  realitate: 1900

după adunare: 0

|model output - reality|

- modelul dă 1
  realitate: 1000

loss de $(1 - 1000)^2 \approx 1\ milion$

- modelul dă 999
  realitate 1000

$$(1000 - 999)^2 = 1^2 = 1$$

- $L[a,b]$

Derivata în funcție de a

Derivata în funcție de b

... Ce se întâmplă dacă urmărim ...

$+$	Crestem pe a $\Rightarrow$ va creste costul
$-$	Descrestem pe a $\Rightarrow$ va creste costul
0	

$\Rightarrow$ gradient mare

minimum local

minimul global

$y = ax + b$

$x = 9 \Rightarrow y = a$
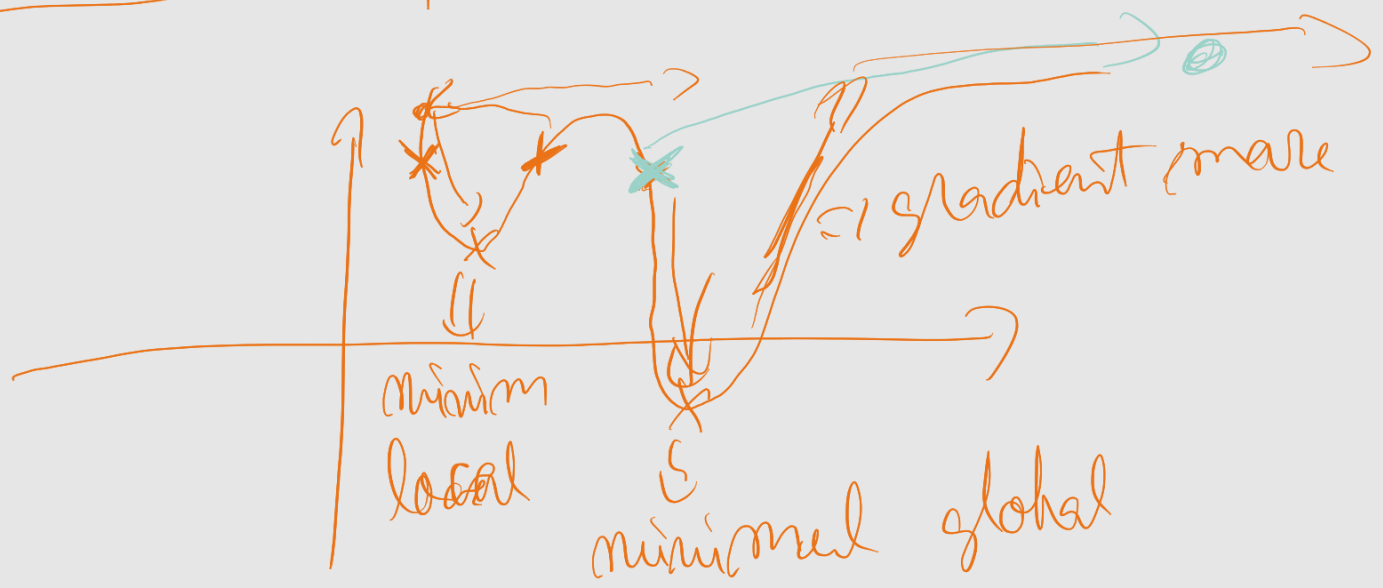
$x = 0 \Rightarrow y = b$

$$\begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1,28} \\ & & & \\ \vdots & P_{28,2} & \cdots & P_{28,28} \\ P_{28,1} & & & \end{pmatrix}$$

$(P_{11}, P_{12}, \cdots P_{1,28}, P_{2,1}, P_{2,2} \cdots P_{2,28} \cdots$

$a_{11} \cdot P_{11} + a_{12} \cdot P_{12} + \ldots + P_{28,28} \cdot a_{28,28} + b = y)$

profit $= a \cdot$ population $+ b$

input pentru ~~unele~~ ~~output~~ cu ~~input~~ ~~ideu~~ de pixeli
$(p_1, p_2 \dots p_{784})$

primul output, $Q_1 = a_1 \cdot p_1 + a_2 \cdot p_2 + \dots + a_{784} \cdot p_{784} + b$

al doilea output, $Q_2 = a_1^{(2)} \cdot p_1 + a_2^{(2)} \cdot p_2 + \dots + a_{784}^{(2)} \cdot p_{784} + b^{(2)}$

$-0.5 \Rightarrow -50\%$

$-1.5 \Rightarrow 1.5 = 150\%$

-10.5

Softmax $[10, 16, 28]$

o modelul ne dă $[1, 2, 3]$
$\qquad 1 \qquad 3$

$e^x > 0$

$$\frac{e^1}{e^1 + e^2 + e^3} \mid \frac{e^2}{e^1 + e^2 + e^3} \mid \frac{e^3}{e^1 + e^2 + e^3}$$

$\underset{0.09 \Rightarrow 1\%}{\underline{\shortparallel}} \qquad \underset{0.24 \Rightarrow 24\%}{\underline{4}} \qquad 0.65 \Rightarrow 65\%$

$$\frac{e^1 + e^2 + e^3}{e^1 + e^2 + e^3} = 1$$

clasa căreia
apartine imaginea
$\downarrow$
pullover

clasa reală	Ce a prezis modelul	Pisatia
t-shirt ⓪	(0.2, 0.3, 0.5)	$-\ln(0.2) = 1.6 \Rightarrow ++$
trousers 1	(0, 1, 0)	$-\ln(1) = 0 \Rightarrow perfect$
dren 2	(0.1, 0, 0.9)	$-\ln(0.9) = 0.1 \Rightarrow greșeală mică$
t-shirt ⓪	(0.5, 0.2, 0.3)	$-\ln(0.5) \approx 0.7 \Rightarrow$

Cross entropy - loss

$y = a \cdot x + b$

$x = 2 \qquad \Rightarrow y = 3 \Rightarrow 300\%$

$a = 1 \qquad 0.3 \Rightarrow 30\%$

$b = 1$

$-\ln(p)$

t = shirt, trouser, shirt

[0.5, -0.7, 1400] $\rightarrow$ shirt

output pt t-shirt $\longrightarrow$ output pt trouser

$\ln(1400) \Rightarrow 10$

$\ln(0.09901) \Rightarrow -\inf$

$\ln(-0.9) \Rightarrow$ max

10×1

$\downarrow$

population și număr de parcuri

10×2

Pt imaginii

batch de 10 imagini, fiecare 28×28

$10 \times 28 \times 28$

$1 \times 28 \times 28$

[1]