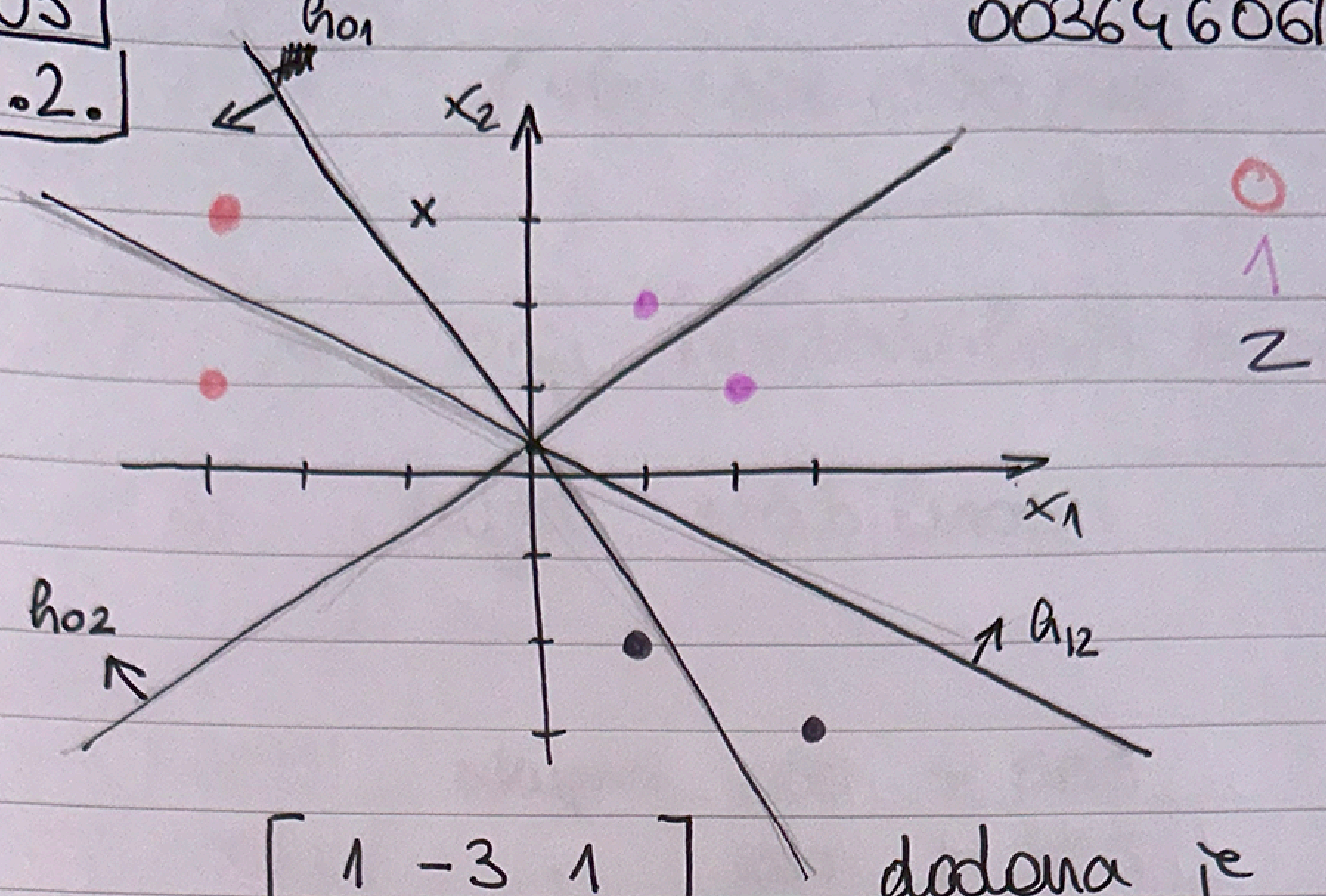
TENA ŠKALEC
0036460611

$\boxed{\text{VO5}}$

$\boxed{1.2.}$



$$\phi(\vec{x}) = \begin{bmatrix} 1 & -3 & 1 \\ 1 & -3 & 3 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & -2 \\ 1 & 2 & -3 \end{bmatrix}$$

dodana je dummy značajka

$\phi(\vec{x})$ ista za sva tri modela

vektori oznaka :

$y_0 = [1\ 1\ 0\ 0\ 0\ 0]^T$

$y_1 = [0\ 0\ 1\ 1\ 0\ 0]^T$

$y_2 = [0\ 0\ 0\ 0\ 1\ 1]^T$

poanta pristupa je klasificirati primjere iz „svoje" klase kao ispravne a sve ostale kao neispravne $\Rightarrow$ 3 binarna klasifikatora

TENA ŠKALEC
0036460611

$\vec{w} = (\phi^T \phi)^{-1} \phi^T \vec{y}$

$\vec{w}_0 = [0{,}33507853 \quad -0{,}21727749 \quad -0{,}0052356]^T$
$\vec{w}_1 = [0{,}2591623 \quad 0{,}23429319 \quad 0{,}22251309]^T$
$\vec{w}_2 = [0{,}40575916 \quad -0{,}01701571 \quad -0{,}21727749]^T$

$h_0(\vec{x}; \vec{w}) = \vec{w}_0^T \vec{x}$
$= 0{,}3351 - 0{,}2173\, x_1 - 0{,}0052\, x_2$

$h_1(\vec{x}; \vec{w}) = 0{,}26 + 0{,}2343\, x_1 + 0{,}2225\, x_2$
$h_2(\vec{x}; w) = 0{,}4058 + (-0{,}017\, x_1) - 0{,}2173\, x_2$

b) $h_{01}(\vec{x}) = h_0(\vec{x}) - h_1(\vec{x})$
$= 0{,}0759 - 0{,}4516\, x_1 - 0{,}2277\, x_2$
$h_{02}(\vec{x}) = h_0(\vec{x}) - h_2(\vec{x})$
$= -0{,}07068 - 0{,}2003\, x_1 + 0{,}21204\, x_2$
$h_{12}(\vec{x}) = -0{,}1466 + 0{,}2513\, x_1 + 0{,}4398\, x_2$

eksplicitni oblici pravaca za skicu:

h_{02}: $x_2 = 0{,}9446\, x_1 + 0{,}3333$
h_{01}: $x_2 = -1{,}9822\, x_1 + 0{,}3333$
h_{12}: $x_2 = -0{,}5715\, x_1 + 0{,}3333$

TENA ŠKALIĆ
0036460611

c) primjer $(-1,3)$

h_{01} stavlja ga u klasu 1
h_{02} u klasu 0
h_{12} u klasu 1

$\Rightarrow$ primjer je klasificiran u klasu 1

d) ova vrsta klasifikacije ne daje probabilistički izlaz, tako da ne možemo iz ovakvog modela dobiti vjerojatnost pripadanja klasi

e) prednost OVR:
manji broj klasifikatora
$\Rightarrow$ manje resursa

nedostatak OVR:
neuravnoteženost broja primjera među klasama

f) problem je funkcija kvadratnog gubitka koja kažnjava pozitivno klasificirane primjere duboko u pozitivnom području, što može dovesti do takvog nagiba decizijske granice da primjeri bliže njoj završe na krivoj strani; npr da dodamo $((1,-100),2)$

TENA SKALIC
0036460611

$\boxed{\text{VO5 - 2.3.}}$

$N = 1000$ / 400 / 300 / 200 / 100
$m = 555$
$K = 4$

OVO, OVR - broj nestabilnih modela

$N > m$ uvjet stabilnosti

OVR

400 - 600	ukupno	1000	> 555
300 - 700	:	1000	> 555
200 - 800	:	1000	> 555
100 - 900		1000	> 555

Svi modeli su stabilni u OVR shemi

OVO

400-300	ukupno	700	> 555
400 - 200	:	600	> 555
400 - 100	:	500	< 555
300 - 200		500	< 555
300 - 100		400	< 555
200 - 100		300	< 555

NESTABILNI

$\boxed{\text{C}}$ 4

$\boxed{\text{V06 - 2.7.}}$

TENA SKALEC
0036460611

$\lambda = 1000 \quad \eta = 0{,}01$

$\phi(\vec{x}) = (1, x_1, x_2, x_1 x_2) = [1 \ \ -1 \ \ 2 \ \ -2]$

$y = 1$

$\vec{w} = [0{,}2 \ \ 0{,}5 \ \ -1{,}1 \ \ 2{,}7]$

$\nabla_w L(y, h(\vec{x})) = (h(\vec{x}) - y)\,\phi(\vec{x})$

$h(\vec{x}) = \sigma(\vec{w}^T \phi(\vec{x}))$

$= \sigma\left([0{,}2 \ \ 0{,}5 \ \ -1{,}1 \ \ 2{,}7] \begin{bmatrix} 1 \\ -1 \\ 2 \\ -2 \end{bmatrix}\right)$

$= \sigma(-7{,}9)$

$= 0{,}0003706$

$h(\vec{x}) - y \approx -1$

$\Rightarrow \nabla_w L(y, h(\vec{x})) \approx [-1 \ \ 1 \ \ -2 \ \ 2]$

za w_1:

$w_1 = w_1(1 - \lambda\eta) - \eta \nabla_w L[1]$

$w_1 = -4{,}51$

promjena težine:

$-4{,}51 - 0{,}5 = -5{,}01$

odgovor: $\boxed{B}$ -5

TENA SKALEC
0036460611

$\boxed{\text{V07-1.1.}}$

a) $\text{softmax}_k(x_1, \ldots, x_n) = \frac{\exp(x_k)}{\sum_j \exp(x_j)}$

$\text{softmax}(2,8,1,5) = ?$

$\sum_j \exp(x_j) = e^2 + e^8 + e^1 + e^5 = 3139{,}478$

$\text{softmax}(2,8,1,5) = [e^2 \; e^8 \; e^1 \; e^5] \div 3139{,}478$
$= [0{,}0024 \quad 0{,}9495 \quad 0{,}0009 \quad 0{,}0473]$

softmax:
1) normalizira vrijednosti tako da zbroj bude 1
2) pojačava veće vrijednosti i smanjuje manje

b) $h_k(\vec{x}, W) = \frac{\exp(\vec{w}_k^T \phi(\vec{x}))}{\sum_j \exp(\vec{w}_j^T \phi(\vec{x}))} = P(y=k \mid \vec{x}, W)$

skup modela $\{h_k\}_k$ daje model multinomijalne logističke regresije

MULTINOMIJALNA LOG. REG

model: $h_k(\vec{x}; W) = \dfrac{\exp(\vec{w}^T \phi(\vec{x}))}{\sum_j \exp(\vec{w}_j \phi(\vec{x}))}$

$= P(y = k \mid \vec{x}, W)$

$W = (\vec{w}_1, \ldots, \vec{w}_K)$

FUNKCIJA POGREŠKE

→ prije smo maksimizirali vjerojatnost oznaka, Bernoulli

⇒ KATEGORIČKA / MULTINOMIJALNA SLUČAJNA V. za K klasa

$\vec{y} = (y_1, \ldots, y_K)^T$ ONE-HOT VECTOR

$\sum_k y_k = 1$ (obviously, Darling)

$\vec{\mu} = (\mu_1, \mu_1 \ldots \mu_K)$ vjerojatnosti pripadnosti klasama

$$\boxed{P(y \mid \mu) = \mu^y (1-\mu)^{1-y}}$$ → BERNOULLI

$$\boxed{P(\vec{y} \mid \vec{\mu}) = \prod_{k=1}^{K} \mu_k^{y_k}}$$

TENA ŠKALEC
0036460611
V07 – 1.1 – c)

$\vec{y} = [0 \quad 0 \quad 1 \quad 0]$

$\vec{\mu} = [0{,}2 \quad 0{,}3 \quad 0{,}4 \quad 0{,}1]$

$\prod_{k=1}^{K} \mu_k^{y_k} = 1 \cdot 1 \cdot 0{,}4 \cdot 1 = 0{,}4$ maska

⇒ doći do f. gubitka / pogreške

$\ln P(\vec{y} \mid X, D) = \ln \prod_{i=1}^{N} P(y^i \mid \vec{x}^i) = \ln \prod_{i=1}^{N} \prod_{k=1}^{K} \mu_k^{y_k^i}$

odakle ovo? od hipoteze

$= \ln \prod_{i=1}^{N} \prod_{k=1}^{K} h_k(\vec{x}^i \mid W)^{y_k^i} = \sum_{i=1}^{N} \sum_{k=1}^{K} y_k^i \ln h(\vec{x}^i \mid W)$

to je izglednost!

$$E(W \mid D) = - \sum_{i=1}^{N} \sum_{k=1}^{K} y_k^i \ln h(\vec{x}^i \mid W)$$

f. gubitka

želimo da izlaz bude što bliže 1 jer $\ln 1 = 0$ dakle mali gubitak

za jako male izlaze, $\ln 0_+$ daje velike neg. brojeve ⇒ ispred je minus ⇒ veliki gubitak