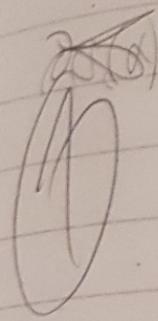


6. LR

4. DZ



$$h(x, w) = \sigma(w^T \phi(x)) = \frac{1}{1 + e^{-w^T \phi(x)}}$$

a) Šir názima rezultate linijstega na $[0, 1]$. To uam
je pogodno za klasifikacija.

b) CROSS-ENTROPY

$$L(h(x), y) = \begin{cases} -\ln h(x), & y=1 \\ -\ln(1-h(x)), & y=0 \end{cases}$$

$$\mathcal{L} = -y \ln h(x) - (1-y) \ln(1-h(x))$$

IZVOD:

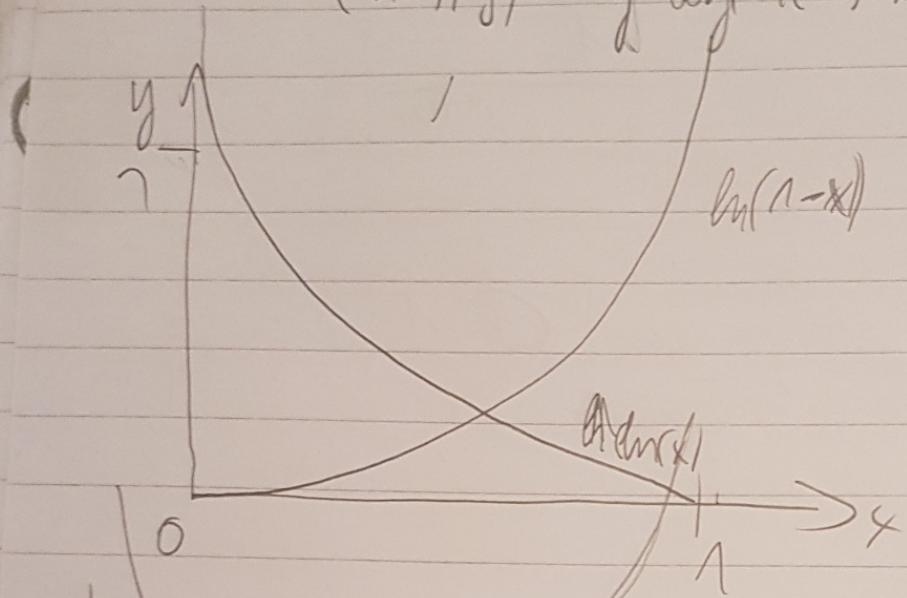
$$\hat{P}(y|x) = h(x)^y (1-h(x))^{1-y}$$

$$\ln L(w|D) = \ln P(D|w) = \ln \prod_{i=1}^N P(y_i|x_i) = \ln \prod_{i=1}^N h(x_i)^{y_i} (1-h(x_i))^{1-y_i}$$

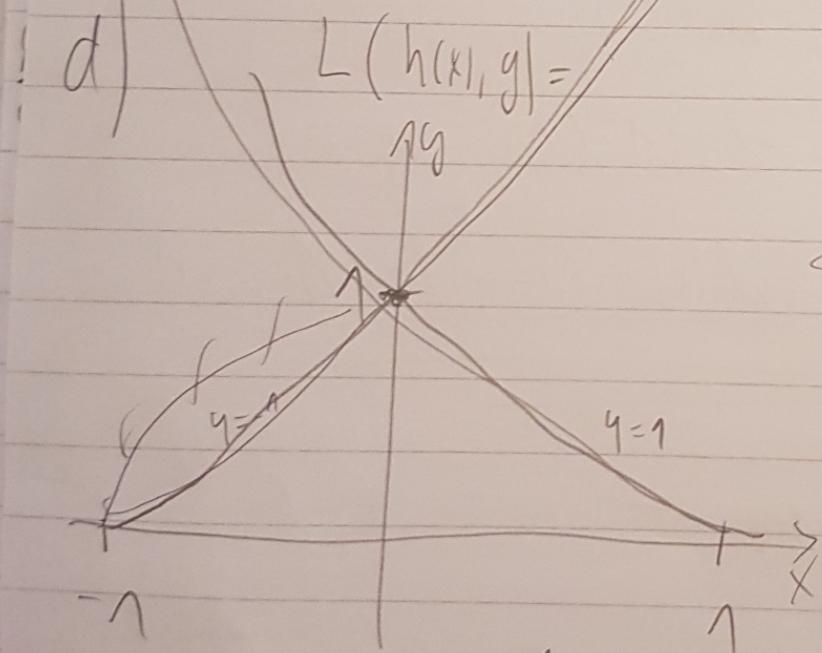
$$E(w|D) = -\ln L(w|D) = -\sum_{i=1}^N \left\{ y_i \ln h(x_i) + (1-y_i) \ln(1-h(x_i)) \right\}$$

$$c) \text{ mit } L(x|w) = y$$

$$L(h(x), y) = -y \log h(x) - (1-y) \log(1-h(x))$$



\rightarrow MAX: ∞
MIN: 0



$$\begin{aligned} g(x) &= -\ln(ax+b) \\ -\ln(ax+b) &= 0 \\ x(b) &= \end{aligned}$$

$$\log_e(ax+b) = 0 \quad \log_{10}(100)$$

$$L(h(x), y) = \frac{1}{\ln 2} \ln(1 + \exp(-y w^T x))$$

$$e^0 = a+b$$

Q.

a) $f: \mathbb{R}^n \rightarrow \mathbb{R}$ konveksna.

akko:

1) Njezina domena je konveksni skup

2) za maki $x_1, x_2 \in \text{dom}(f)$ i maki $\alpha \in [0, 1]$:

$$f(x) = f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha)f(x_2)$$

b) fja je unimodala.

$$f(\alpha x + (1-\alpha)y) \leq \max \{f(x), f(y)\}$$

$$\alpha f(x) + (1-\alpha) \cancel{f(y)} \leq \max \{f(x), f(y)\}$$

a) Tvorivo minimum výroby f.j. Sedan od něj je
práv správ. Odvozeno z něj počtem těch (typ. intervalek) i
~~počtem~~ prvního a u třetího zadílu na dílnu
kterého se číslo bude blíže min. Velikou neg. horizont odvozenou

3.

a)

$$\nabla E(w, D) = \left(\partial_w h(x_i) \right) x_i$$

$$\text{GRADIENT} = \sum_{i=1}^N (h(x_i) - y_i) x_i$$

$$E(w, D) = \frac{1}{N} \sum_{i=1}^N (-y_i \ln h(x_i) - (1-y_i) \ln(1-h(x_i)))$$

(POSSIBLE ENTRY ERROR?)

GRADIENTNI SPUT:

BATCH:

$$w \leftarrow (0, 0, \dots, 0)$$

Ponavljaj do konvergencije

$$\Delta w \leftarrow (0, \dots, 0)$$

$$\forall i=1 \dots N \quad \phi(w_i x_i)$$

$$\Delta w \leftarrow \Delta w - (h - y_i) x_i$$

η \leftarrow optimalni kri. param u myern - Δw

$$w \leftarrow w - \eta \Delta w$$

STOASTICKI:

$$w \leftarrow (0, 0, \dots, 0)$$

ponavljaj do homogenije

permutaciju D

$$x_i \quad i=1 \dots N$$

$$h \leftarrow \sigma(w^T x_i)$$

$$\Delta w \leftarrow \Delta w + (h - y_i) x_i$$

Monte Carlo pred u vrijeme $-t w$

$$w \leftarrow w - \eta \Delta w$$

d) brže, ali iterativno

He je slup
u svakom koraku

MS

8

a)

$$h_k(x, w) = \frac{\exp(w_k^T \phi(x))}{\sum_j \exp(w_j^T \phi(x))} = P(y=k | x, w)$$

$$w = (w_1, \dots, w_n)$$

SOFTMAX

$$f(\text{softmax}) = \frac{\exp(w_k^T \phi(x))}{\sum_j \exp(w_j^T \phi(x))}$$

$$\varphi(y|\mu) = \prod_{k=1}^K y_k^{y_k}$$

$$\ln P = \ln \prod_i^N \varphi(y_i|x_i) = \ln \prod_i^N \prod_{k=1}^K \mu_k^{y_k} = \ln \prod_i^N \prod_{k=1}^K h_k(x_i|w)$$
$$= \sum_i^N \sum_{k=1}^K y_k^i \ln h_k(x_i|w)$$

$$\mathbb{E}(w|D) = - \sum_i^N \sum_{k=1}^K y_k^i \ln h_k(x_i|w)$$