

7.

→ REGRESIJA

a)

$$X = \mathbb{R}$$

$$h(x) = w_1 x + w_0$$

$$E(h|D) = \frac{1}{2} \sum_{i=1}^N (y^{(i)} - (w_1 x^{(i)} + w_0))^2$$

→ želimo h koji minimizira pogrešku

$$\frac{\partial}{\partial w_0} = \frac{1}{2} \cdot 2 \sum_{i=1}^N (-y^{(i)} + (w_1 x^{(i)} + w_0))$$

$$0 = \sum_{i=1}^N (-y^{(i)} + w_1 x^{(i)} + w_0)$$

$$\sum_{i=1}^N -y^{(i)} + \sum_{i=1}^N w_1 x^{(i)} + N w_0 = 0$$

$$w_0 = \frac{\sum_{i=1}^N w_1 x^{(i)}}{N} + \frac{\sum_{i=1}^N y^{(i)}}{N}$$

$$w_0 = w_1 \bar{x} + \bar{y}$$

$$\frac{\partial}{\partial w_1} = \frac{1}{2} \cdot 2 \sum_{i=1}^N (-y^{(i)} + w_1 x^{(i)} + w_0) \cdot x^{(i)}$$

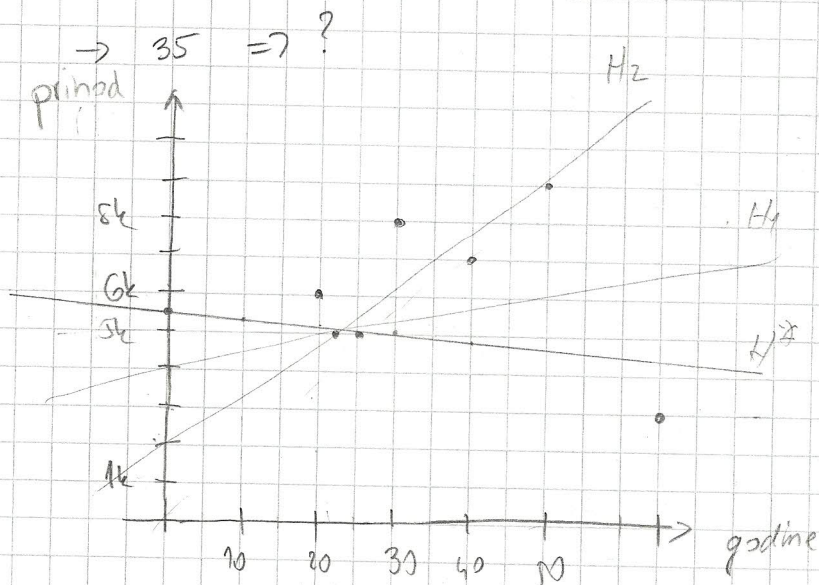
$$\sum_{i=1}^N (-y^{(i)} x^{(i)} + w_1 (x^{(i)})^2 + w_0 x^{(i)}) = 0$$

$$\sum_{i=1}^N x^{(i)} y^{(i)} - w_1 \sum_{i=1}^N (x^{(i)})^2 - (w_1 \bar{x} + \bar{y}) \cdot N \bar{x} = 0$$

$$\sum_{i=1}^N x^{(i)} y^{(i)} - w_1 \sum_{i=1}^N (x^{(i)})^2 - N w_1 \bar{x}^2 - N \bar{x} \bar{y} = 0$$

$$w_1 = \frac{\sum_{i=1}^N x^{(i)} y^{(i)} - N \bar{x} \bar{y}}{\sum_{i=1}^N (x^{(i)})^2 - N \bar{x}^2}$$

b) $D = \{(20, 6k), (22, 5k), (25, 5k), (30, 8k), (40, 7k), (50, 9k), (65, 3k)\}$



$$y = w_1 x_1 + w_0$$

\Rightarrow naučna hipoteza

$$h(x) = -0.016x + 5.564$$

$$w_0 = w_1 \bar{x} + \bar{y}$$

$$w_1 = \frac{\sum_{i=1}^N x_i y_i - N \bar{x} \bar{y}}{\sum_{i=1}^N x_i^2 - N \bar{x}^2}$$

$$\bar{x} = \frac{20+22+25+30+40+50+65}{7} = 36$$

$$\bar{y} = \frac{6+5+5+8+7+9+3}{7} = 6.14$$

$$\sum_{i=1}^N x_i y_i = 20 \cdot 6 + 22 \cdot 5 + 25 \cdot 5 + 30 \cdot 8 + 40 \cdot 7 + 50 \cdot 9 + 65 \cdot 3 = 1520$$

$$\sum_{i=1}^N x_i^2 = 20^2 + 22^2 + 25^2 + 30^2 + 40^2 + 50^2 + 65^2 = 10734$$

$$w_1 = \frac{1520 - 7 \cdot 36 \cdot 6.14}{10734 - 7 \cdot 36^2} = -0.016$$

$$w_0 = -0.016 \cdot 36 + 6.14 = 5.564$$

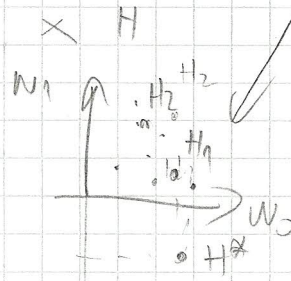
$$h(35) = 6.124 \Rightarrow 6k$$

c) $E(h_2|D) > E(h_1|D) > E(h^*|D) \leftarrow E(h^*|D) = 12.92211$

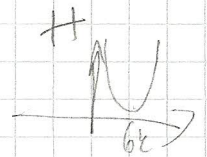
$$E(h|D) = \frac{1}{2} \sum_{i=1}^N (w_1 x^i + w_0 - y^i)^2$$

ulazni parametri \Rightarrow

prostor hip $\rightarrow 2$ p.k.



svaka hip. je
točka u



d) složeniji model bio bi prikladniji za ovaj problem jer bi se bolje prilagodio ulaznim primjerima
s obzirom na izgled u. pod.

8. a) 3 osnovne komponente algoritma strojnog učenja:

1. MODEL (prostor hipoteza)

2. FUNKCIJA GUBITKA

3. OPTIMIZACIJSKI POSTUPAK

b) Pristranost ograničavanjem (pristranost jerika)

\hookrightarrow MODEL

Pristranost preferencije (p. pretraživanja)

\hookrightarrow FUNK. G. i OPT. POSTUPAK

$$\begin{matrix} E(h^*|D) \\ = E_0[L] \end{matrix}$$

c) LINEARNA REGRESIJA

$$1) h(\vec{x}) = \sum_{i=1}^N w_i x^{(i)} + w_0$$

$$2) L(y^{(i)}, h(x^{(i)}|\theta)) = (h(w; x^{(i)} + w_0) - y^{(i)})^2$$

$$3) \theta^* = \underset{\theta}{\operatorname{argmin}} E(\theta|D)$$

\rightarrow metoda najmanjeg kvadrata

$$E[L] = \int p(\theta) L(\theta) d\theta$$

na skupu za učenje \leftarrow empirijska pog.

9. a) Funkcija gubitka

$$L(y^{(i)}, h(x^{(i)}|\theta)) = (h(x^{(i)}|\theta) - y^{(i)})^2$$

$$E(\theta|D) = \frac{1}{N} \sum_{i=1}^N L(y^{(i)}, h(x^{(i)}|\theta))$$

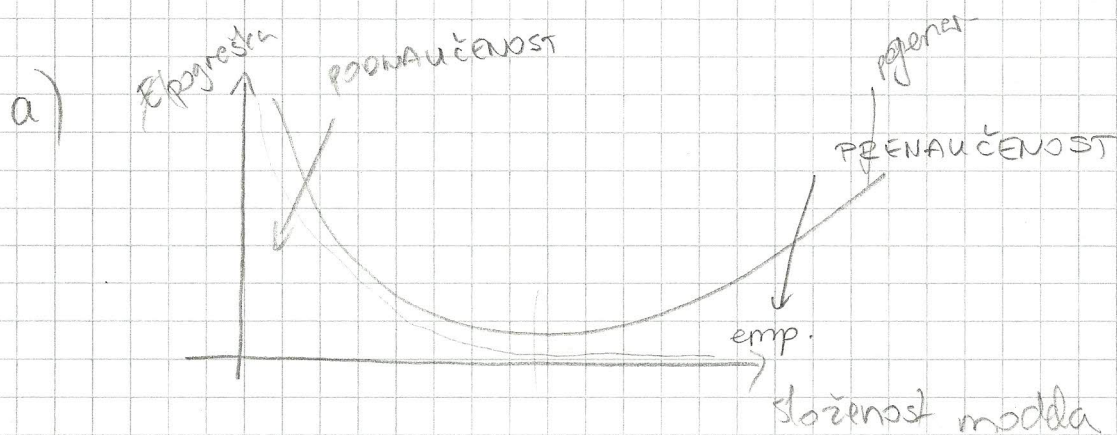
b) Matrica gubitka za neželjene e-poštu

	spam	! spam
spam	0	5
! spam	2	0

↑
veća
penaliz.

$$E(\theta|D) = \frac{1}{7} \cdot [5 \cdot (1-0) + 2 \cdot (0-1)] = \frac{3}{7}$$

10.



b) H_α , α - hiperparametar

pogreška gener. > empirijska p.

=> model koji smo izabrali je prenaučen
tj. bolje evaluira primj. za učenje
(previše im se prilagodio) i stoga

probleme što
nema
gube mo

c) H_α

već $\alpha \Rightarrow$ složeniji model

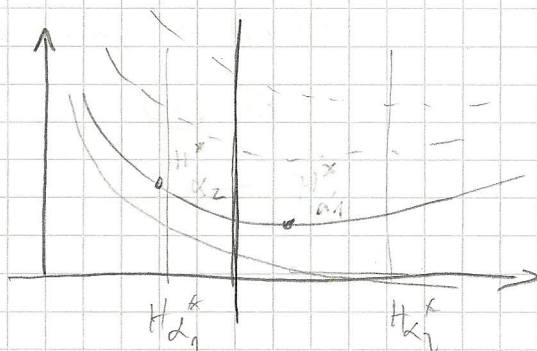
optim. alg. L_1 i L_2

$$E(\theta_2 | D) > E(\theta_1 | D)$$

$$L_1 \rightarrow \alpha_1^*$$

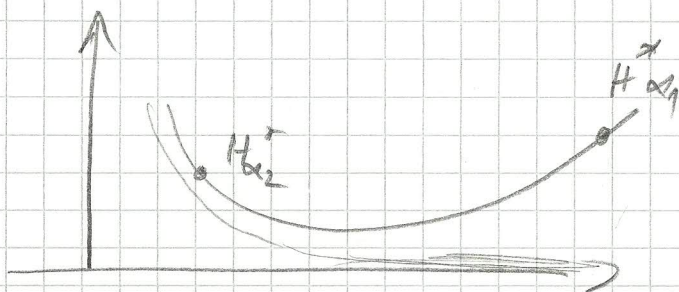
$$L_2 \rightarrow \alpha_2^*$$

$\Rightarrow L_1$ daje bolje
rez. na primj.
za učenje



--- L_2
— L_1

d) Model učen lošijim alg. L_2 može imati
manju grešku od modela učenim alg. L_1
zato što može biti prenaučten



odabrati hiperparam: naći odabrati model

