Postupak učenja decizijskih funkcija potencijalnim funkcijama

Dosadašnji postupci → unaprijed smo pretpostavili oblik decizijske funkcije

Potencijalne funkcije dopuštaju učenje decizijskih funkcija bez pretpostavke o konačnom obliku decizijske funkcije

Decizijske funkcije mogu se generirati uporabom potencijalnih funkcija na temelju uzoraka $\vec{x}_k$, $k=1,2,\ldots$,
Potencijalna funkcija za neki uzorak $\vec{x}$ može se opisati kao:

funkcije Tipa 1

$$K(\vec{x},\vec{x}_k)=\sum_{i=1}^{\infty}\lambda_i^2\varphi_i(\vec{x})\varphi_i(\vec{x}_k)$$

gdje su $\varphi_i(\vec{x})$ $i=1,2,\ldots$ ortonormalne funkcije i λ_i realni brojevi $\neq 0$ izabrani tako da je $K(\vec{x},\vec{x})$ ograničena za $\vec{X}_k \in \omega_1 \cup \omega_2$.

Decizijska funkcija $d(\vec{x})$ konstruira se iz sekvence potencijalnih funkcija $K(\vec{x}, \vec{x}_1)$, $K(\vec{x}, \vec{x}_2), \ldots$ koje odgovaraju slijedu uzoraka za učenje $\vec{x}_1, \vec{x}_2, \ldots$ i koji su predstavljeni stroju tijekom procesa učenja.

Potencijalne funkcije:

Kontinuirane funkcije koje zadovoljavaju slijedeće uvjete:

1. $K(\vec{x}, \vec{x}_k) = K(\vec{x}_k, \vec{x})$
   - simetrična
2. $K(\vec{x}, \vec{x}_k)$ ima maksimum pri $\vec{x} = \vec{x}_k$
3. $K(\vec{x}, \vec{x}_k) \rightarrow 0$ ako udaljenost $\|\vec{x} - \vec{x}_k\| \rightarrow \infty$

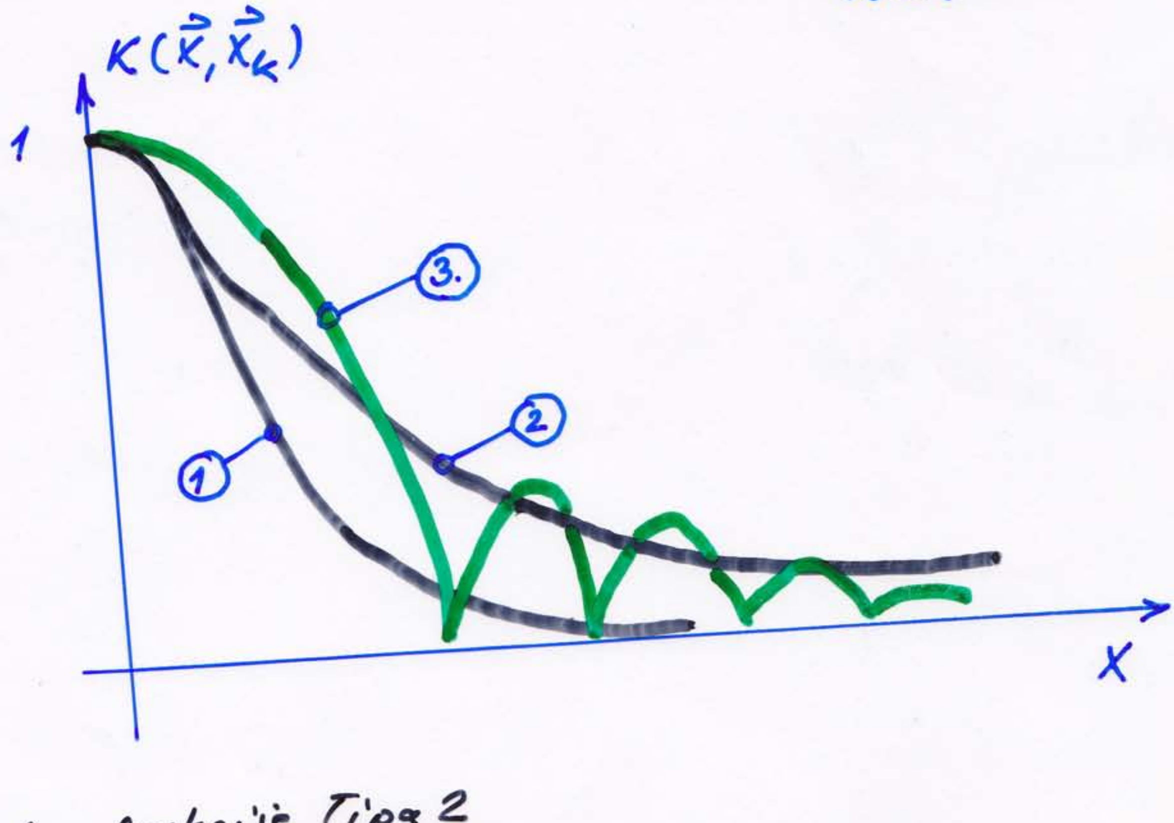
Primjeri potencijalnih funkcija:

1. $K(\vec{x}, \vec{x}_k) = e^{-\alpha \|\vec{x} - \vec{x}_k\|^2}$; $\alpha > 0$

2. $K(\vec{x}, \vec{x}_k) = \dfrac{1}{1 + \alpha \|\vec{x} - \vec{x}_k\|^2}$

3. $K(\vec{x}, \vec{x}_k) = \left| \dfrac{\sin \alpha \|\vec{x} - \vec{x}_k\|^2}{\alpha \|\vec{x} - \vec{x}_k\|^2} \right|$

Ilustracija potencijalnih funkcija za $n = 1$



Potencijalne funkcije Tipa 2

Osnovna zamisao:

- Potencijalne funkcije opisuju potencijal točke $\vec{x}$ (u prostoru značajki) zbog naboja u $\vec{x}_k$ (u točki prostora)
- Ako je potencijal točke $\vec{x}$ zbog jediničnog naboja u točki $\vec{x}_k$ jednak $K(\vec{x}, \vec{x}_k)$ onda je u toj točki prosječni potencijal zbog jediničnih naboja N_i uzoraka iz razreda ω_i jednak

$$\frac{1}{N_i} \sum_{k=1}^{N_i} K(\vec{x}, \vec{x}_k)$$

- Uzorci iz razreda ω_1 i ω_2
- Uzorci su u prostoru predočeni točkama kojima je dodijeljen jedinični naboj

uzorcima iz $\omega_1 \rightarrow$ pozitivan jedinični naboj

uzorcima iz $\omega_2 \rightarrow$ negativan jedinični naboj

$+q$ svakom uzorku $\vec{x}_k \in \omega_1$

$-q$ svakom uzorku $\vec{x}_k \in \omega_2$

$k = 1, 2, \ldots, N$

- Uzorci zbog svog naboja prouzrokuju potencijalno polje - potencijal

- Potencijal ima vršnu vrijednost u samoj točki $\vec{X}_k$ i pada s udaljenosti od $\vec{X}_k$

- Uzorci koji pripadaju razredu ω_1 oblikuju "pozitivni plato" s točkama koje su smještene na vrhovima tog platoa

- Uzorci iz ω_2 oblikuju "negativni plato"

- Dva platoa zbog različitosti naboja ($+q$ i $-q$) oblikuju dva polja koja su odijeljena udolinom u kojoj je potencijal jednak nuli

- Udolina → granica između razreda → decizijska funkcija

Postupak učenja:

N uzoraka u skupu za učenje

$\vec{X}_i \in \omega_1 \cup \omega_2$

Pretpostavimo da se decizijska funkcija oblikuje ovako:

$$d_k(\vec{x}) = d_{k-1}(\vec{x}) + r_k K(\vec{x}, \vec{x}_k), \text{ gdje je}$$

$$d_0(\vec{x}) = 0$$

r_k je korekcijski faktor

$$r_k = \begin{cases} 0 & \text{ako je } \vec{X}_k \in \omega_1 \text{ i } d_{k-1}(\vec{X}_k) > 0 \\ 0 & \text{ako je } \vec{X}_k \in \omega_2 \text{ i } d_{k-1}(\vec{X}_k) < 0 \\ 1 & \text{ako je } \vec{X}_k \in \omega_1 \text{ i } d_{k-1}(\vec{X}_k) \leq 0 \\ -1 & \text{ako je } \vec{X}_k \in \omega_2 \text{ i } d_{k-1}(\vec{X}_k) \geq 0 \end{cases}$$

za $k=1$ (uzimamo prvi uzorak iz skupa za učenje)

$$d_1(\vec{X}_1) = 0 + K(\vec{x}, \vec{x}_1)$$
ako je $\vec{X}_1$ iz ω_1

$$d_1(\vec{X}_1) = 0 - K(\vec{X}, \vec{X}_1) \text{ ako je}$$
$\vec{X}_1$ iz ω_2

Za k+1 korak:

- Ako je $\vec{X}_{k+1} \in \omega_1$ i

$d_k(\vec{X}_{k+1}) > 0$ ili

$\vec{X}_{k+1} \in \omega_2$ i

$d_k(\vec{X}_{k+1}) < 0$ tada

$$d_{k+1}(\vec{X}) = d_k(\vec{X})$$

- Ako je $\vec{X}_{k+1} \in \omega_1$ i

$d_k(\vec{X}_{k+1}) \leq 0$ onda

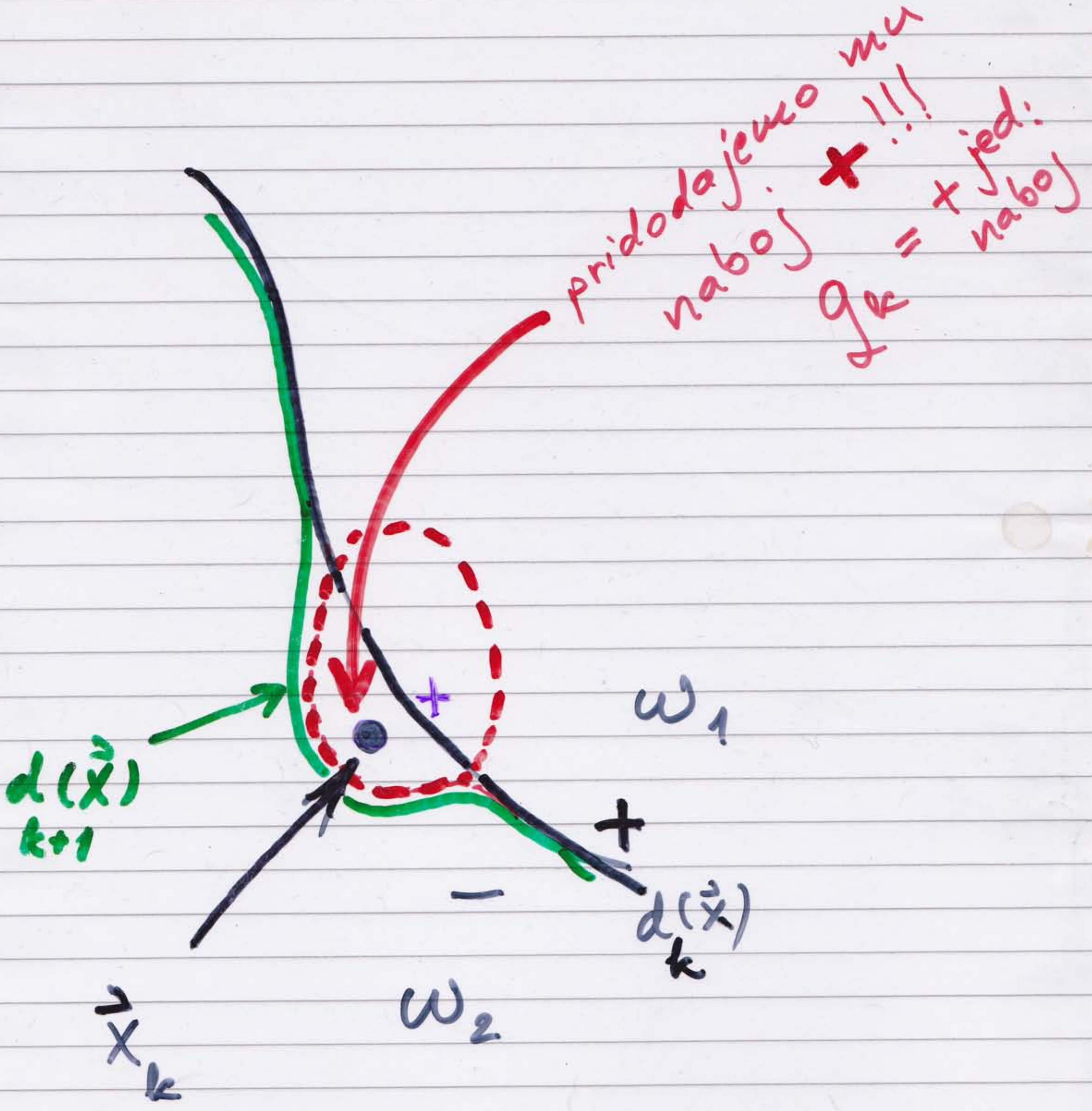
$$d_{k+1}(\vec{X}) = d_k(\vec{X}) + K(\vec{X}, \vec{X}_{k+1})$$

- Ako je $\vec{X}_{k+1} \in \omega_2$ i $d_k(\vec{X}_{k+1}) \geq 0$

onda

$$d_{k+1}(\vec{X}) = d_k(\vec{X}) - K(\vec{X}, \vec{X}_{k+1})$$

Grafička interpretacija:



$\vec{X}_k \in \omega_1$ ima pozitivan naboj ali $d(\vec{x}) = 0$ ga svrstava u ω_2

POVEĆATI NABOJ $\vec{X}_k$ TAKO DA SE $d(\vec{x})$ mijenja!

Poopćenje postupka za $M>2$ razreda:

- razred ω_i:

svaka točka područja $\mathcal{P}_i$ u prostoru značajki koji odgovara uzorcima iz ω_i ima prosječni potencijal (zbog naboja uzoraka iz ω_i) veći od potencijala zbog naboja uzoraka iz ostalih razreda

$$d_i(\vec{x}) = \frac{1}{N_i} \sum_{k=1}^{N_i} K(\vec{x}, \vec{x}_k)$$

W. S. Meisel (1969. godine) je pokazao da se uvijek može naći takva potencijalna funkcija $K(\vec{x}, \vec{x}_k)$ da granicom

$$d(\vec{x}) = d_i(\vec{x}) - d_j(\vec{x}) = 0$$

pravilno razgraničavamo uzorke iz skupa za učenje iz razreda ω_i i ω_j.

Postupak učenja za $M>2$

$$d_0^{1}(\vec{x}),\ d_0^{2}(\vec{x}) \dots d_0^{M}(\vec{x})=0$$

- u k-tom koraku učenja obrađujemo uzorak $\vec{x}_k$ koji pripada razredu ω_i
- decizijsku funkciju u k-tom koraku određujemo na sljedeći način:

• ako je $\vec{x}_k \in \omega_i$

i $d_{k-1}^{(i)}(\vec{x}_k) > d_{k-1}^{(j)}(\vec{x}_k)$ za $\forall j \neq i$

funkciju odlučivanja ne mijenjamo

$$d_k^{(i)}(\vec{x}) = d_{k-1}^{(i)} \quad i=1,2,\dots,M$$

• ako je $\vec{X}_k \in \omega_i$ i za neki l

je $d_{k-1}^{(i)}(\vec{x}_k) \leq d_{k-1}^{(l)}(\vec{x}_k)$

tada MORAMO POPRAVITI DEC. FUNKCIJE:

$$d_k^{(i)}(\vec{x}) = d_{k-1}^{(i)} + K(\vec{x}, \vec{x}_k)$$

$$d_k^{(l)}(\vec{x}) = d_{k-1}^{(l)} - K(\vec{x}, \vec{x}_k)$$

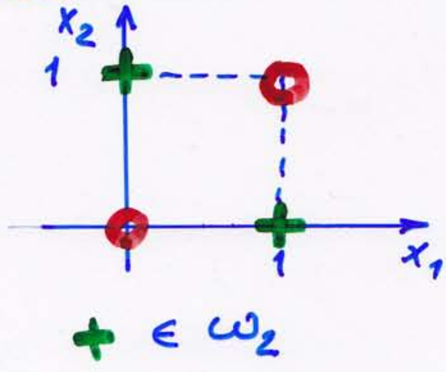
$$d_k^{(j)}(\vec{x}) = d_{k-1}^{(j)}(\vec{x}) \quad j=1,2,\dots,M \quad j\neq i;\ j\neq l$$

Primjer:

Zadana su dva razreda uzoraka:

$\omega_1 = \{(0,0)^T, (1,1)^T\}$

$\omega_2 = \{(0,1)^T, (1,0)^T\}$



Izaberimo potencijalnu funkciju $K(\vec{x}, \vec{x}_k) = e^{-\alpha \|\vec{x} - \vec{x}_k\|^2}$

$\alpha = 1$

Za $n=2$ (dimenzionalnost uzorka) potencijalna funkcija ima oblik:

$$K(\vec{x}, \vec{x}_k) = e^{-[(x_1 - x_{k1})^2 + (x_2 - x_{k2})^2]}$$

Započnimo postupkom učenja:

- $\vec{x}_1 = (0,0)^T \in \omega_1$ prvi uzorak za učenje

$$d_1(\vec{x}) = d_0(\vec{x}) + K(\vec{x}, \vec{x}_1) \; ; \; d_0(\vec{x}) = 0$$

$$d_1(\vec{x}) = e^{-[(x_1-0)^2 + (x_2-0)^2]}$$

$$d_1(\vec{x}) = e^{-(x_1^2 + x_2^2)}$$

- uzorak $\vec{x}_2 = (1,1) \in \omega_1$ vrijedi

$$d_1(\vec{x}_2) = e^{-(1+1)} = e^{-2} > 0$$

zato je: $d_2(\vec{x}) = d_1(\vec{x}) = e^{-(x_1^2 + x_2^2)}$

- za uzorak $\vec{x}_3 = (0,1)^T \in \omega_2$ je

$$d_2(\vec{x}_3) = e^{-(x_1^2 + x_2^2)} = e^{-1} > 0$$

PAZI $\vec{x}_3$ je iz ω_2 !

$$d_3(\vec{x}) = d_2(\vec{x}) - K(\vec{x}, \vec{x}_3)$$

$$d_3(\vec{x}) = e^{-(x_1^2 + x_2^2)} - e^{-[(x_1-0)^2 + (x_2-1)^2]}$$

- za sljedeći uzorak $\vec{x}_4 = (1,0)^T \in \omega_2$

dobivamo:

$$d_3(\vec{x}_4) = e^{-(1^2)} - e^{-[1^2 + (-1)^2]}$$
$$= e^{-1} - e^{-2} > 0$$

Moramo popraviti funkciju odlučivanja

$$d_4(\vec{x}) = d_3(\vec{x}) - K(\vec{x}, \vec{x}_4)$$

$$d_4(\vec{x}) = e^{-(x_1^2 + x_2^2)} - e^{-[x_1^2 + (x_2-1)^2]} - e^{-[(x_1-1)^2 + x_2^2]}$$

Provjeravamo da li se funkcijom $d_4(\vec{x})$ može pravilno razvrstati sve uzorke iz skupa za učenje!

$$\vec{x}_5 = \vec{x}_1 = (0,0)^T \in \omega_1$$

$$d_4(\vec{x}_5) = e^{-0} - e^{-1} - e^{-1} > 0$$

$$d_5(\vec{x}) = d_4(\vec{x})$$

• za $\vec{X}_6 = \vec{X}_2 = (1,1)^T \in \omega_1$

$$d_5(\vec{X}_6) = e^{-2} - e^{-1} - e^{-1} < 0$$

Treba popraviti decizijsku funkciju!

$$d_6(\vec{x}) = d_4(\vec{x}) + K(\vec{x}, \vec{X}_6)$$

$$d_6(\vec{x}) = e^{-(x_1^2 + x_2^2)} - e^{-[x_1^2 + (x_2-1)^2]} - e^{-[(x_1-1)^2 + x_2^2]} + e^{-[(x_1-1)^2 + (x_2-1)^2]}$$

• za $\vec{X}_7 = \vec{X}_3 = (0,1)^T \in \omega_2$ je

$d_6(\vec{X}_7) = e^{-1} - e^{0} - e^{-2} + e^{-1} < 0$ zato je

$$d_7(\vec{x}) = d_6(\vec{x})$$

• za $\vec{X}_8 = \vec{X}_4 = (1,0)^T \in \omega_2$ je

$d_7(\vec{X}_8) = e^{-1} - e^{-2} - e^{0} + e^{-1} < 0$ zato je

$$d_8(\vec{x}) = d_7(\vec{x})$$

• za $\vec{X}_9 = \vec{X}_1 = (0,0)^T \in \omega_1$ je

$d_8(\vec{X}_9) = e^{0} - e^{-1} - e^{-1} + e^{-2} > 0$ zato je

$$d_9(\vec{x}) = d_8(\vec{x})$$

- za $\vec{X}_{10} = \vec{X}_2 = (1,1)^T \in \omega_1$

$$d_9(\vec{X}_{10}) = e^{-2} - e^{-1} - e^{-1} + e^{0} > 0$$

$$d_{10}(\vec{X}) = d_9(\vec{X})$$

Sve smo uzorke iz skupa za učenje pravilno razvrstali funkcijom odlučivanja:

$$d(\vec{X}) = e^{-(x_1^2+x_2^2)} - e^{-[x_1^2+(x_2-1)^2]} - e^{-[(x_1-1)^2+x_2^2]} + e^{-[(x_1-1)^2+(x_2-1)^2]}$$

Funkcija odlučivanja ima onoliko eksponencijalnih članova koliko je bilo KOREKCIJA u postupku učenja.