1.3. ŠEST POSTULATA RU

Postulat 1: U cilju prikupljanja informacija o RU moraju biti raspoloživi reprezentativni uzorci iz M razreda.

Postulat 2: Jednostavan uzorak ima značajke koje karakteriziraju njegovu pripadnost određenom razredu.

Postulat 3: Značajke uzoraka koji pripadaju jednom razredu uzoraka zauzimaju kompaktno područje u prostoru značajki. Područja okupirana značajkama različitih razreda su odvojena.

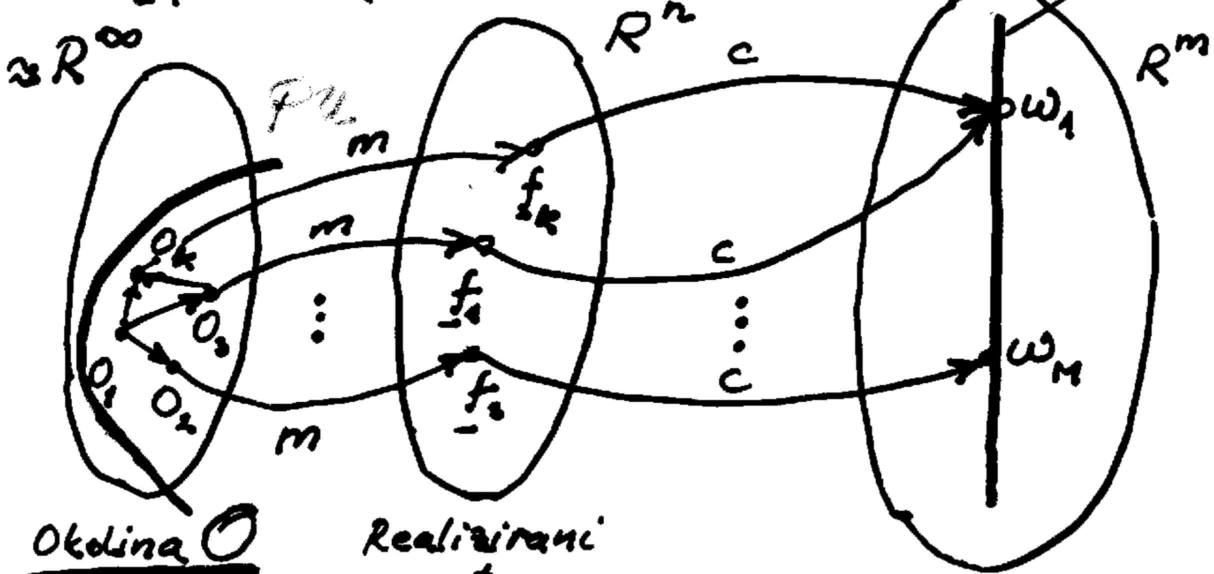
Postulat 4: Složeni (kompleksni) uzorak sastoji se iz jednostavnijih građevnih komponenti ili segmenata objekata koji se nalaze u izvjesnim odnosima. Uzorak se može sastaviti iz tih komponenti.

<u>Postulat 5:</u> Složeni uzorak koji pripada određenom području uporabe ima određenu strukturu. To implicira da bilo kakvo uređenje jednostavnih građevnih elemenata neće dati uzorak $\underline{f}_k(\vec{x})$.

<u>Postulat 6:</u> Dva uzorka su si slična ako je pogodno definirana mjera udaljenosti u prostoru značajki mala.

2. MODEL SUSTAVA ZA RASPOZNAVANJE UZORAKA

2.1 FORMALNI MODEL



Okolina $\mathcal{O}$

$o_k \in \mathcal{O}$
↳ objekt
$v_j \in V\mathcal{O}$
↳ veze

Realizirani prostor uzoraka

$\underline{f}_k$ - uzorci

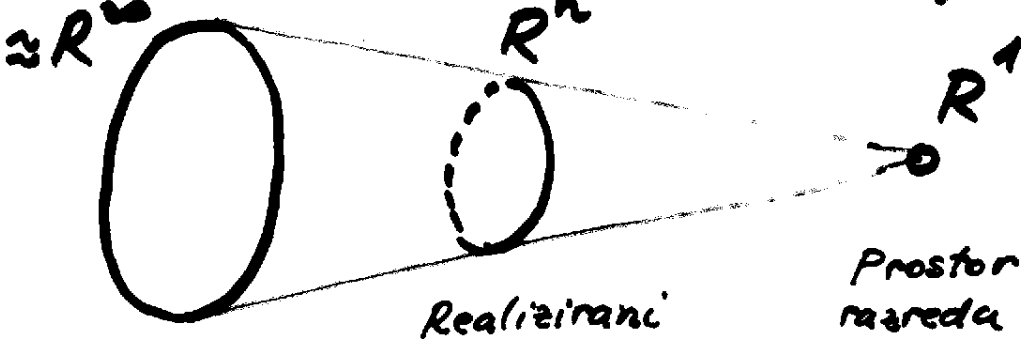
Prostor razreda uzoraka

$\Omega = \{\omega_1, \omega_2, \dots, \omega_M\}$

$$m : o_k \mapsto \underline{f}_k \qquad c : \underline{f}_k \mapsto \omega_i$$

m - mjerenje

Uhrov stožac raspoznavanja - redukcija informacija



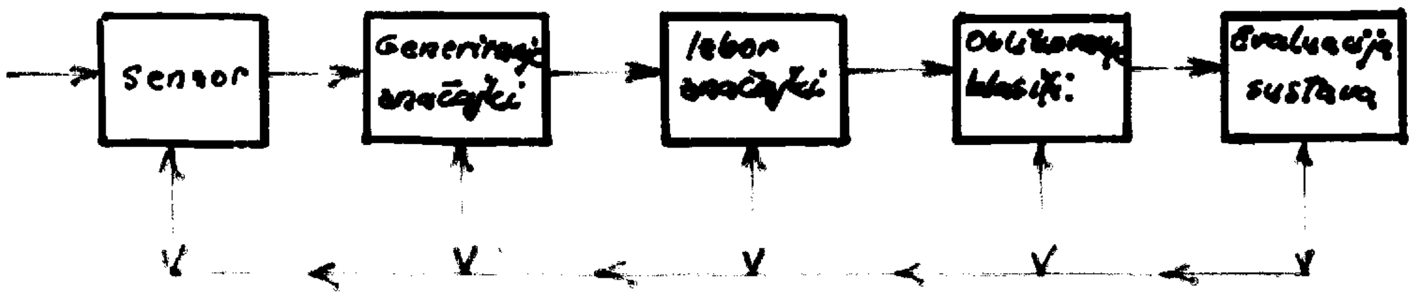
Okolina

Realizirani prostor uzoraka

Prostor razreda uzoraka

2.2. ZNAČAJKE, VEKTOR ZNAČAJKI I KLASIFIKATOR

Osnovne faze/stupnjevi u postupku oblikovanja sustava RU



značajke – promatramo ih kao slučajne varijable : x_i

vektor značajki – slučajni vektori $\vec{x}$

$$\vec{x} = [x_1, x_2, \ldots, x_n]^T$$

- značajke koje predstavljaju razlike između razreda uzoraka nazivaju se <u>INTERSET</u> značajke

<u>INTRASET značajke</u> su zajedničke SVIM razredima iz PU (područja uporabe) i <u>ne nose</u> diskriminacijsku informaciju – TAKVE ZNAČAJKE MOGU SE ZANEMARITI

Izbor značajki – izlučiti i izabrati INTERSET značajke

- U većini slučajeva određivanje potpunog skupa diskriminacijskih značajki je IZNIMNO TEŠKO ILI ČAK NEMOGUĆE;

- Neke diskriminacijske značajke mogu se naći na temelju raspoloživih rezultata mjerenja (senzoriranja);

- Redukcija dimenzionalnosti vektora značajki uporabom transformacija uz minimalni gubitak informacije;

- Vektor značajki predočen kao točka u n-dimenzionalnom prostoru značajki;

- Obično definiramo i neku vrstu metrike u takvom prostoru značajki;

- Klasifikacija (razvrstavanje) uzorka temelji se na decizijskim funkcijama; PROBLEM: određivanje optimalne decizijske procedure;

- Problem klasifikacije može se promatrati kao razvrstavanje nepoznatog uzorka u potprostor prostora značajki na temelju decizijskih granica koje definiraju te potprostore!

- Decizijske granice određene su decizijskim funkcijama:

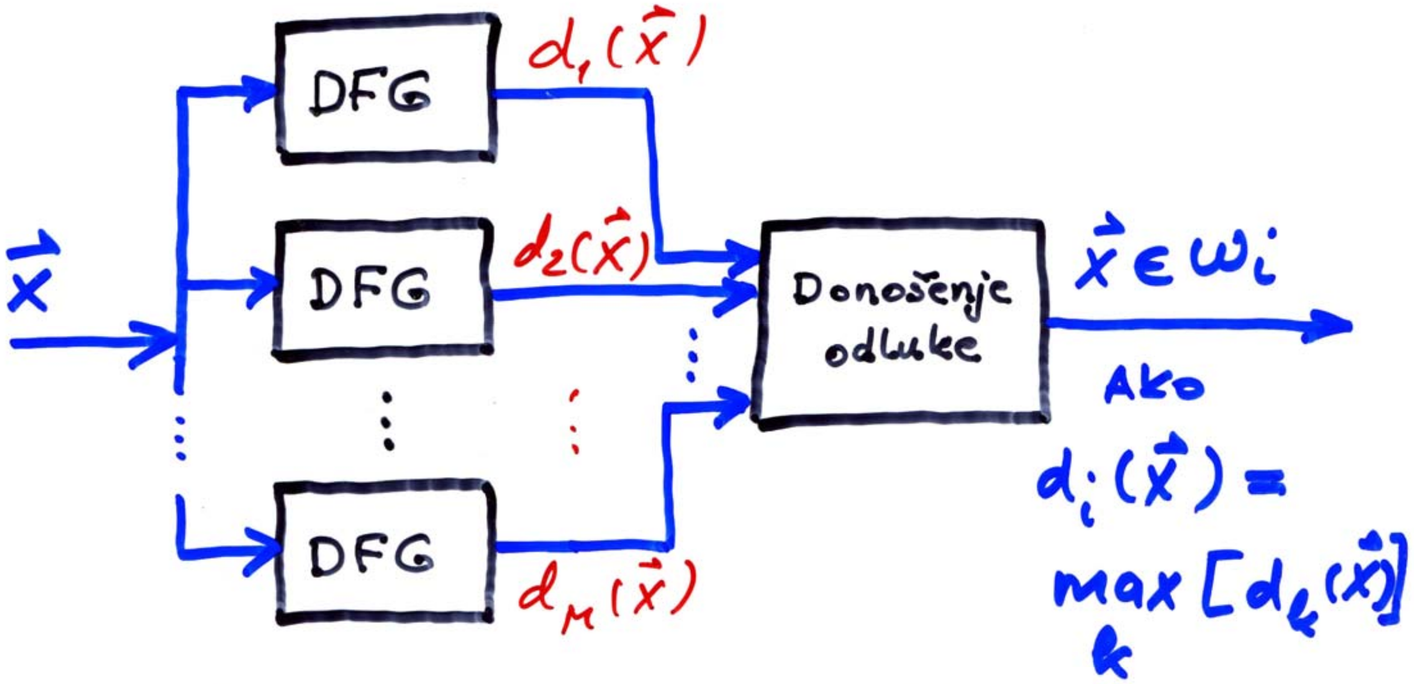
$$d_1(\vec{x}), d_2(\vec{x}), \ldots, d_M(\vec{x})$$

/ d_i je funkcija koja ima za argument VEKTOR a vraća SKALAR / VAŽNO!

- Pravilo razvrstavanja:

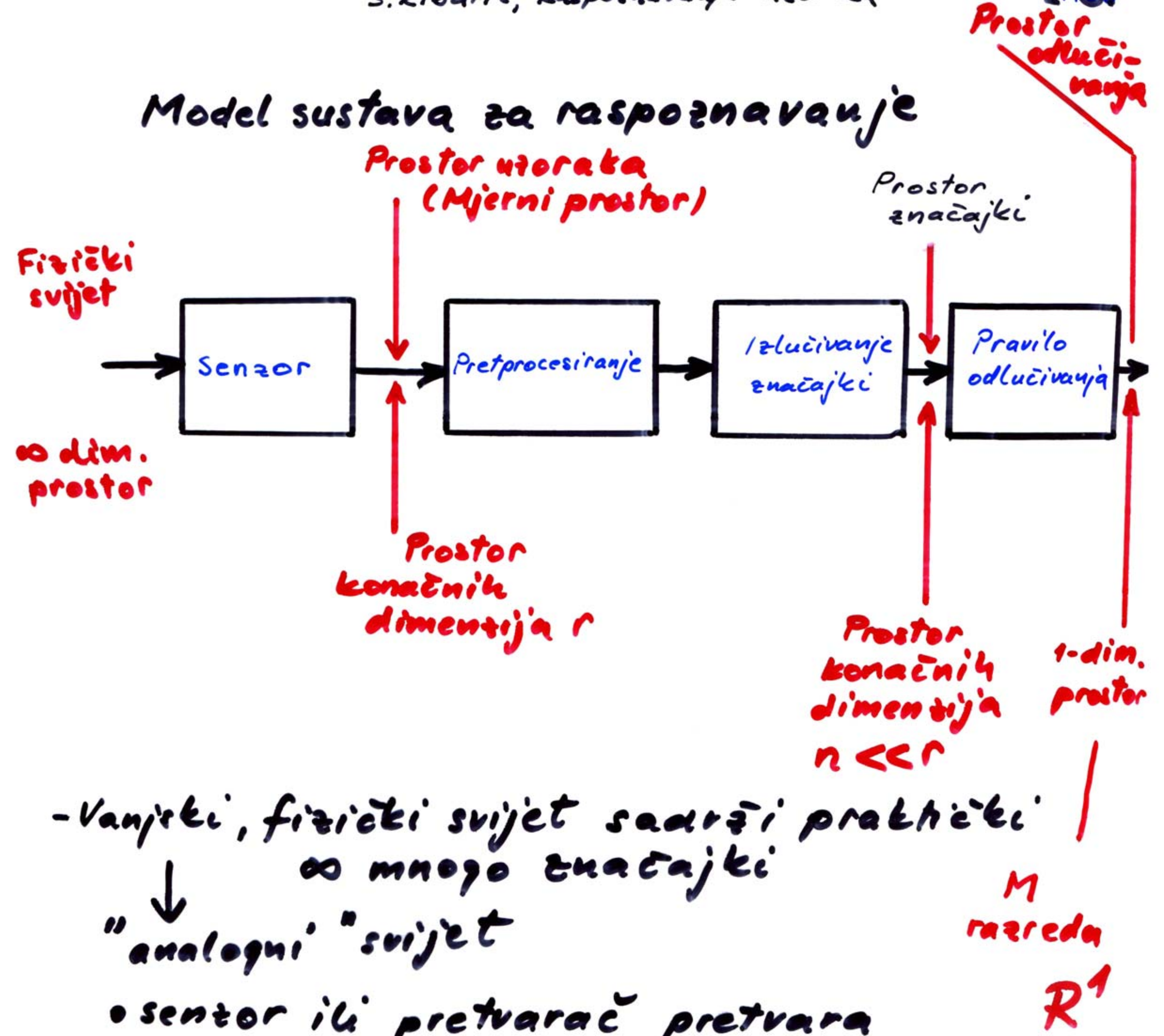
Ako $d_i(\vec{x}) > d_j(\vec{x})$ za $i, j = 1, 2, \ldots, M$ te $j \neq i$ tada nepoznati uzorak $\vec{x}$ pripada razredu ω_i.

Blok-dijagram klasifikatora



DFG - Generator decizijske funkcije (engl. Decision function generator)

Model sustava za raspoznavanje



- Vanjski, fizički svijet sadrži praktički ∞ mnogo značajki

↓

"analogni" svijet

- senzor ili pretvarač pretvara analogni svijet u zapis koji sadrži r (brojčanih) vrijednosti

- Pretprocesiranje: izlučivanje šuma, poboljšanje mjernog podatka
- Prostor značajki: $n << r$

  $\vec{x} = (x_1, x_2, x_3, \ldots, x_n)^T$
- Prostor odlučivanja – jednodim. prostor R^1

PRIMJER

$\mathcal{OB} = \{ O_k ; \; k = 1, 2, \ldots \}$

O_i - signal EKG-a pacijenta

PU - automatska dijagnoza signala EKG-a u testu opterećenja

Uzorak

$$\underline{f}_k(x) \longrightarrow \vec{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Razred objekata:

Normalni EKG
Granični EKG
Nenormalni EKG

Razred uzoraka

$C_1 \,/\, \omega_1$ - Nenormalni EKG
$C_2 \,/\, \omega_2$ - Granični EKG
$C_3 \,/\, \omega_3$ - Normalni EKG

Skup uzoraka za učenje ili vježbanje $\mathcal{U}_M = (S_N, \Omega)$

Tip uzorka - jednostavni uzorak

PRIMJER (nastavak)

Sustav za automatsku dijagnozu EKG-a /Lesterova dijagnostička metoda/

EKG u testu opterećenja

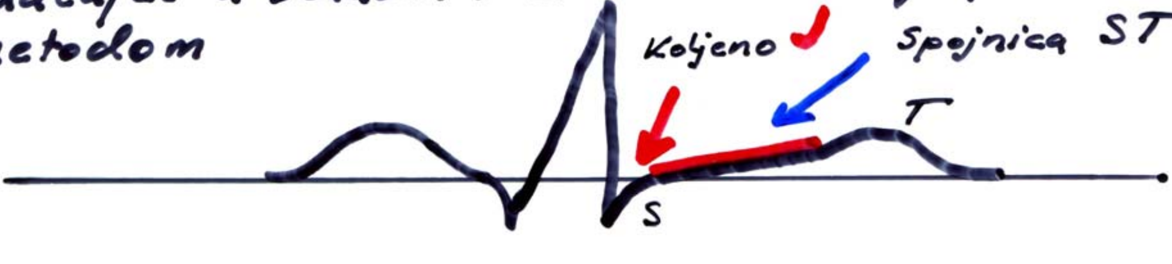
- Vanjski, fizički svijet

- Uzorkovanje (f = 800Hz) (A/D konverzija)

- Pretprocesiranje :
  - Eliminacija ekstrasistole
  - Detekcija stabilnih točaka (maksimalna i minimalna derivacija)
  - Filtriranje usrednjavanjem /eliminacija šuma za faktor $\sqrt{n}$; gdje je n broj analiziranih perioda

Utvrđivanje broja otkucaja u minuti

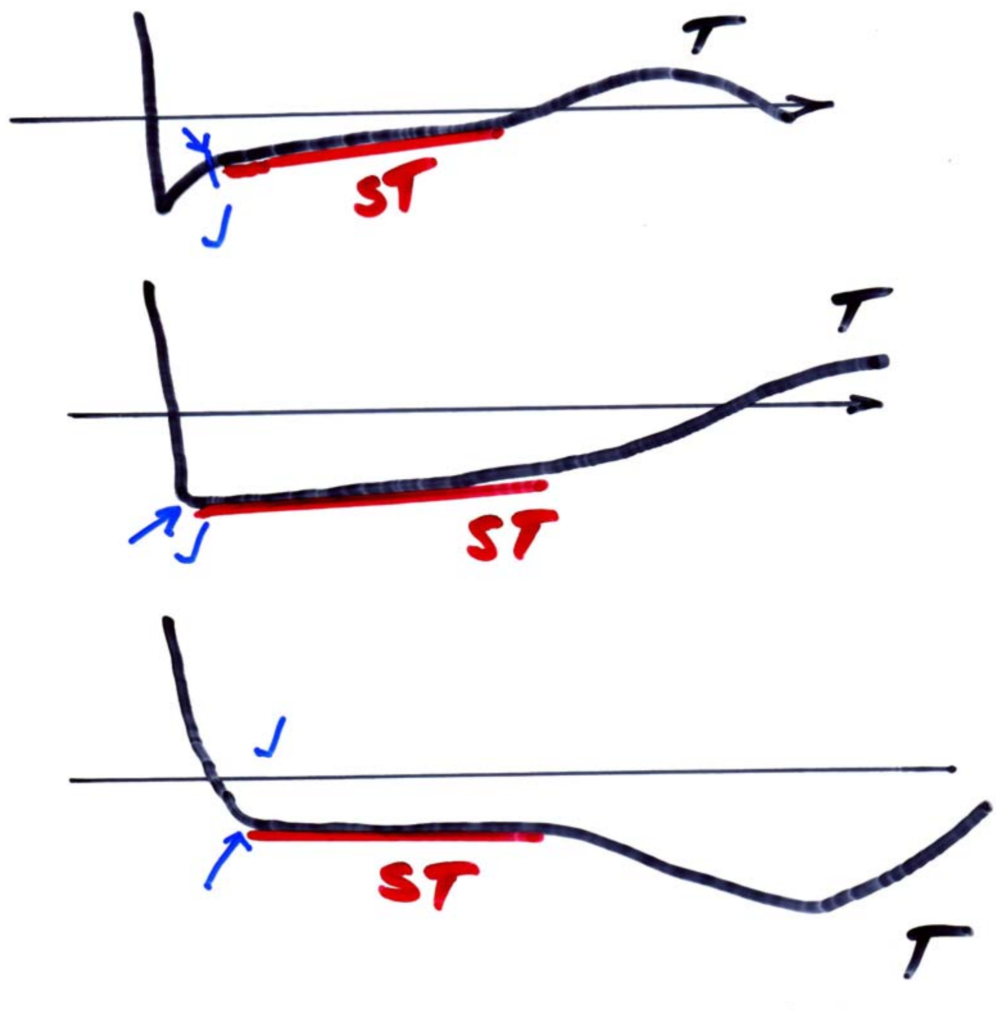
- Izlučivanje značajki u skladu s Lesterovom dijagnostičkom metodom



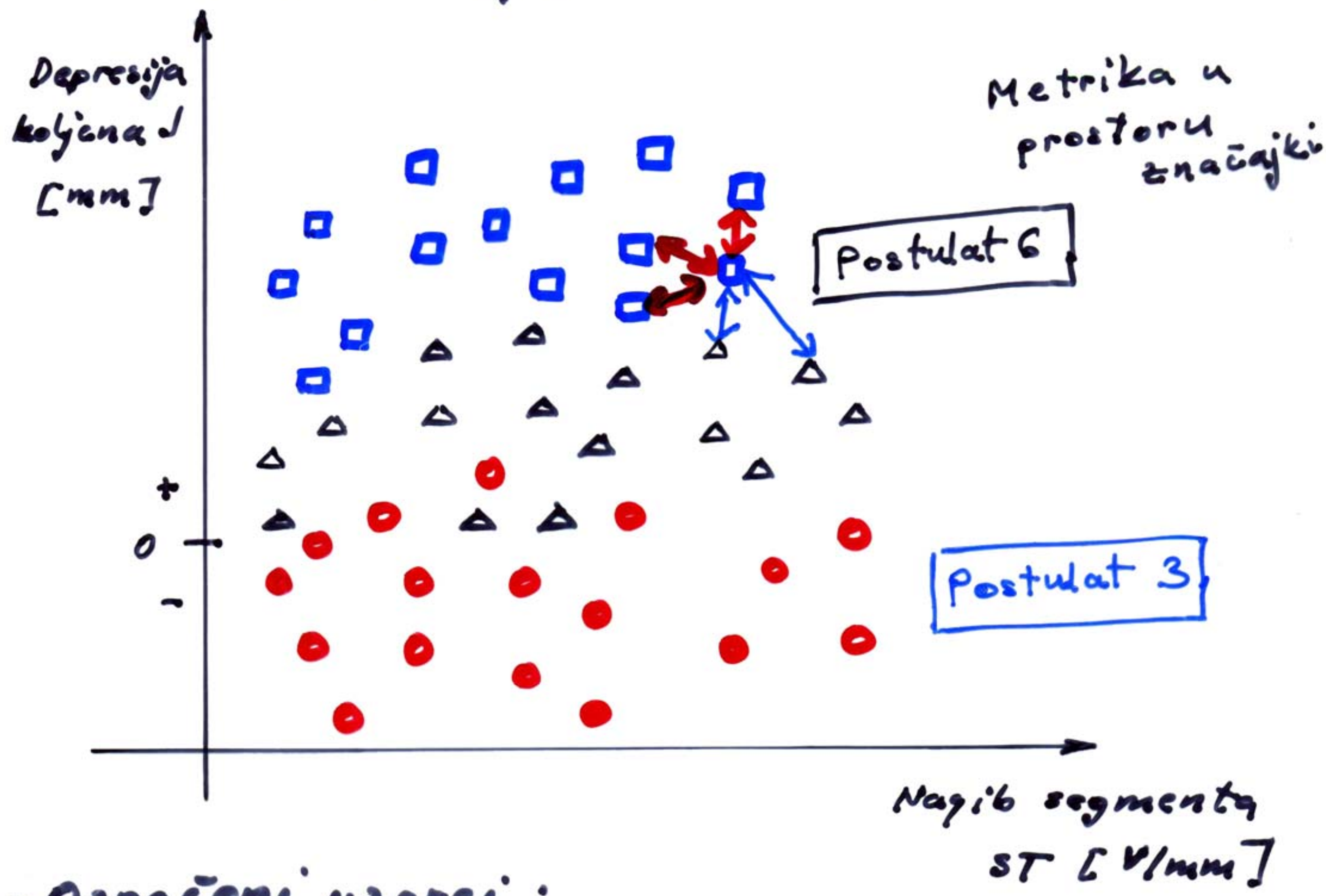
Problem: - koljeno J ima
različite oblike

- segment ST se nalazi
između kraja (završetka) vala
S i početka vala T i
duljina segmenta ST zavisi
od frekvencije signala EKG-a
(broja otkucaja srca u minuti)

Primjeri kliničkih oblika koljena J i
nagiba segmenata ST



Prostor značajki



- Označeni uzorci:
  - uzorci iz skupa za učenje

● – Nenormalan EKG
△ – Granični slučaj
□ – Normalan EKG

u skladu s postulatom 1

$$\Omega = \{\underbrace{\text{Nenormalan EKG}}_{\omega_1}, \overbrace{\text{Granični slučaj}}^{\omega_2}, \underbrace{\text{Normalan EKG}}_{\omega_3}\}$$

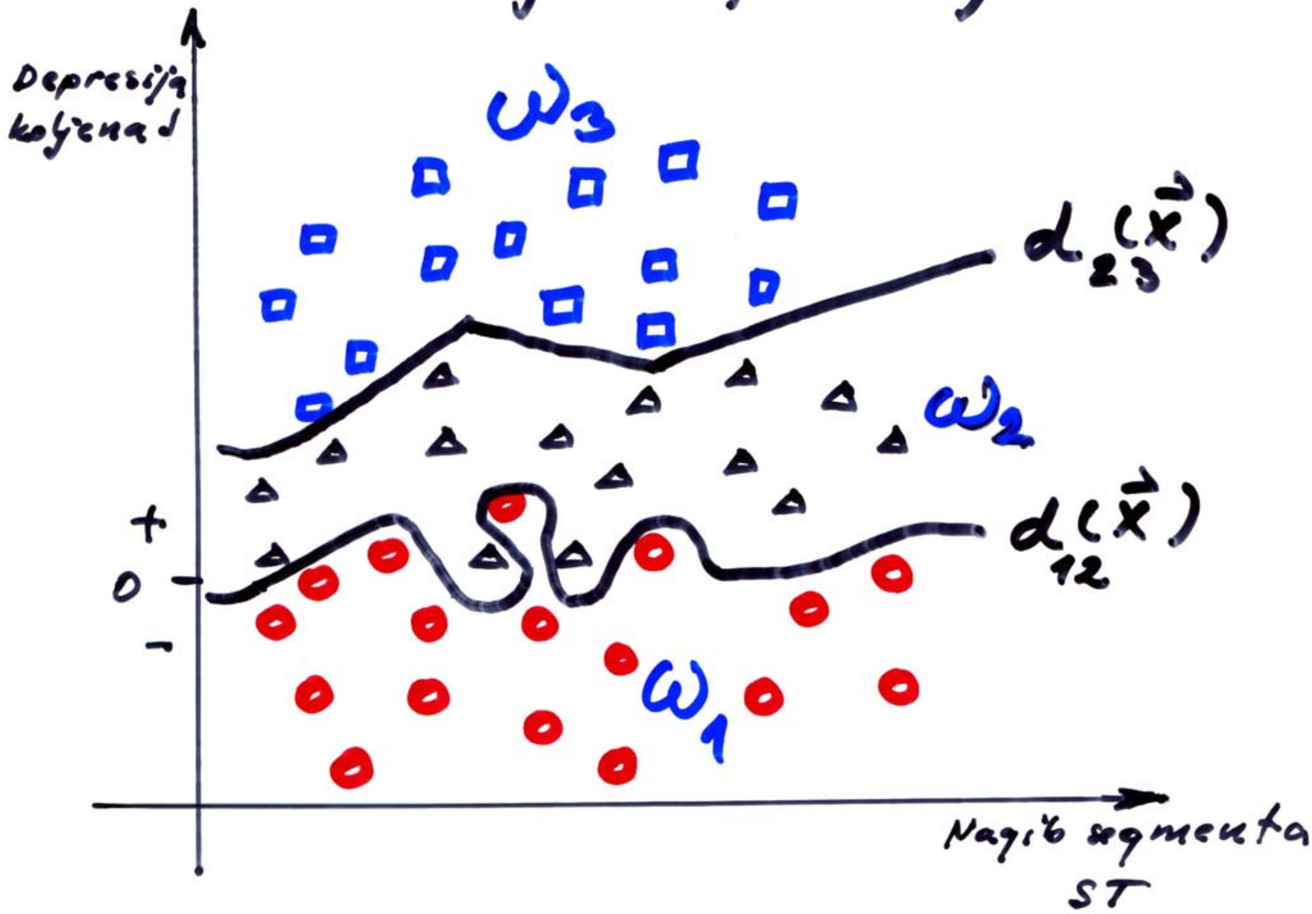
$$\Omega = \{\omega_1, \omega_2, \omega_3\}; \quad M = 3$$

Postulat 2

vektor značajki $\vec{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

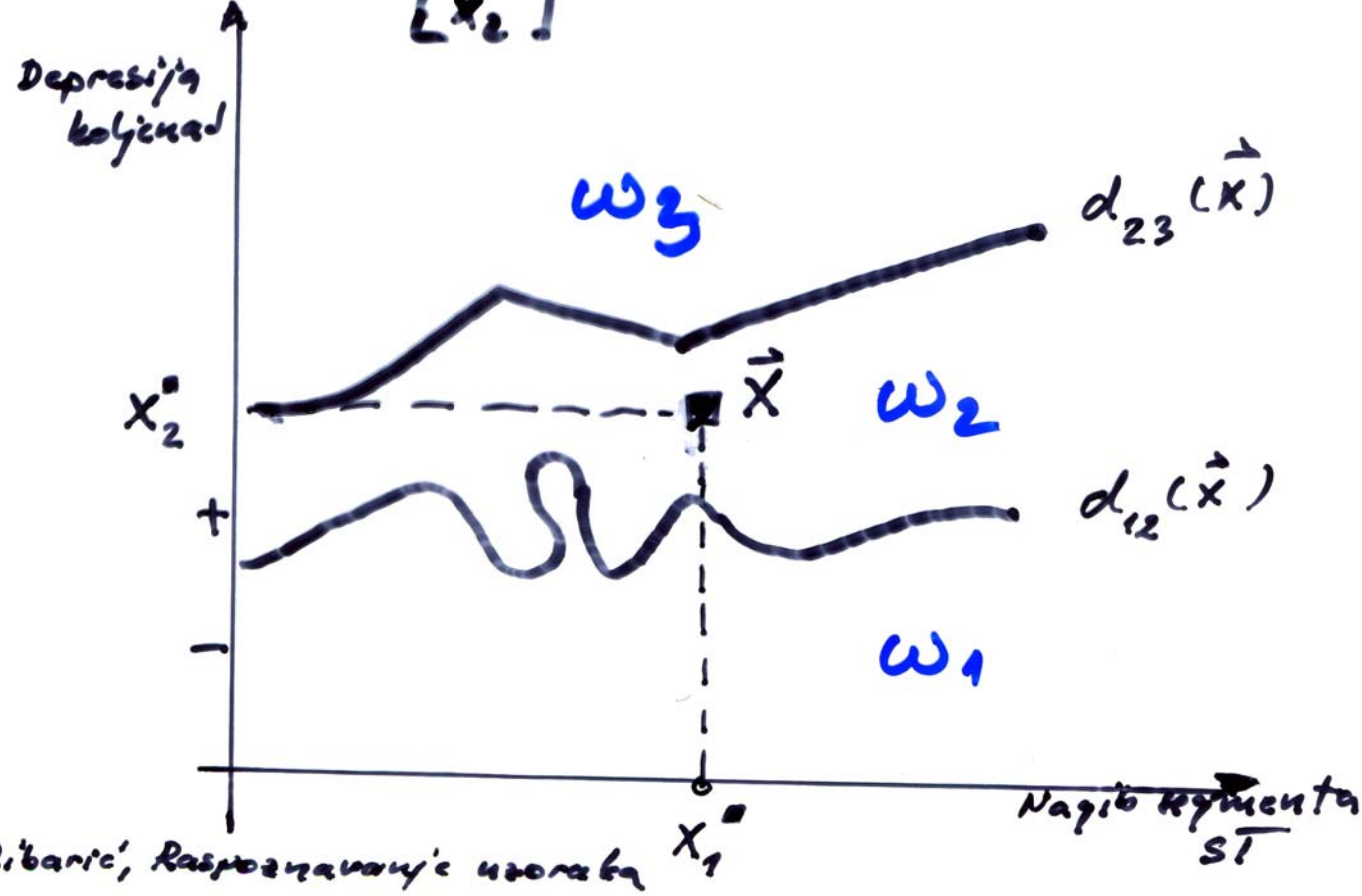
x_1 – nagib segmenta ST
x_2 – depresija koljena J

- decizijske funkcije



Razvrstavanje (klasifikacija) nepoznatog uzorka (■)

$$\vec{x} = \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix}$$



LINEARNE FUNKCIJE ODLUČIVANJA

(engl. Linear discriminant functions)

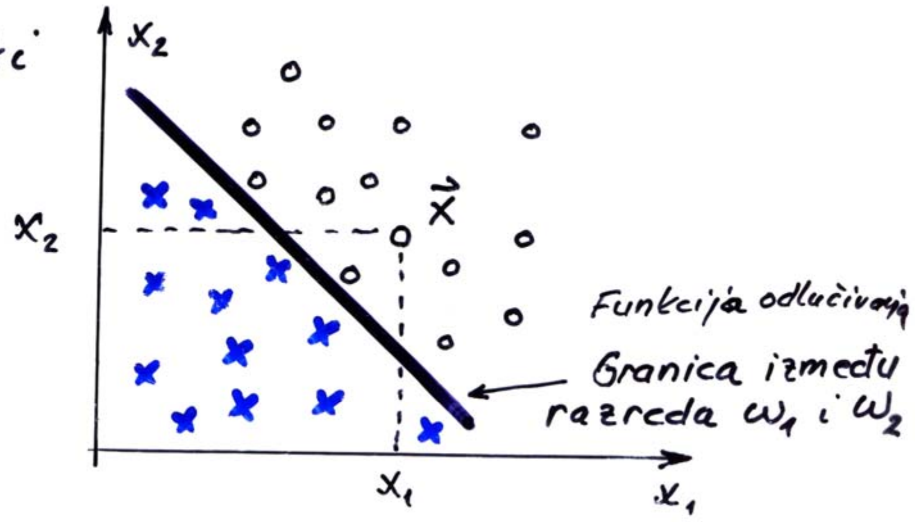
$\vec{X}$ - vektor značajki

$$\vec{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$\circ \in \omega_1$

$\times \in \omega_2$

$M = 2$



Funkcija odlučivanja kao linearna kombinacija komponenti vektora $\vec{X}$:

$$d(\vec{X}) = W_1 x_1 + W_2 x_2 + \dots + W_n x_n + W_{n+1}$$

za

$n = 2$ funkcija odlučivanja $\rightarrow$ jednadžba pravca

$n = 3$ jednadžba ravnine

$n > 3$ hiperravnina

W_i - težinski koeficijenti ; $i = 1, 2, \dots, n$

W_{n+1} - pomaknuće, utežnosni prag (engl. bias, threshold weight)

SLUČAJ: DVA RAZREDA (ω_1 i ω_2)

M=2

$$d(\vec{x}) = W_1 X_1 + W_2 X_2 + \dots + W_n \cdot X_n + W_{n+1} = 0$$

$M=2$; ω_1 i ω_2

Decizijsko pravilo:

ako $d(\vec{x}) > 0$ onda $\vec{x} \in \omega_1$

$d(\vec{x}) < 0$ onda $\vec{x} \in \omega_2$

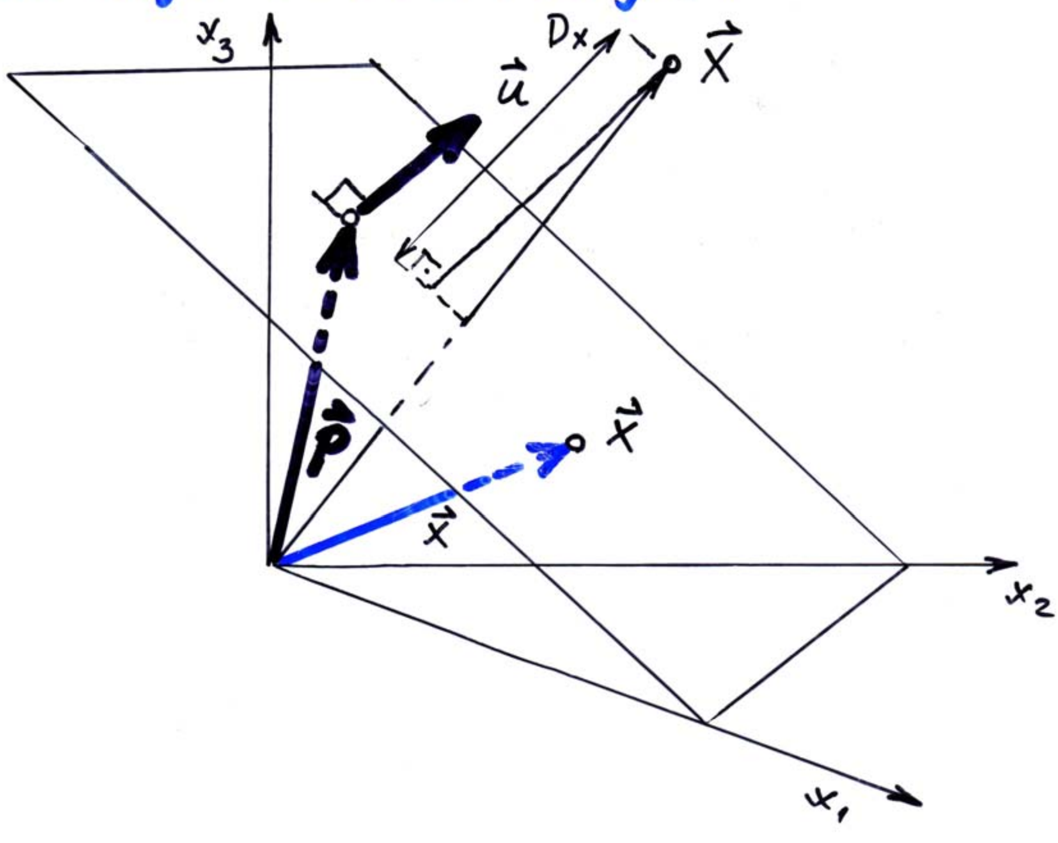
ako je $d(\vec{x}) = 0$

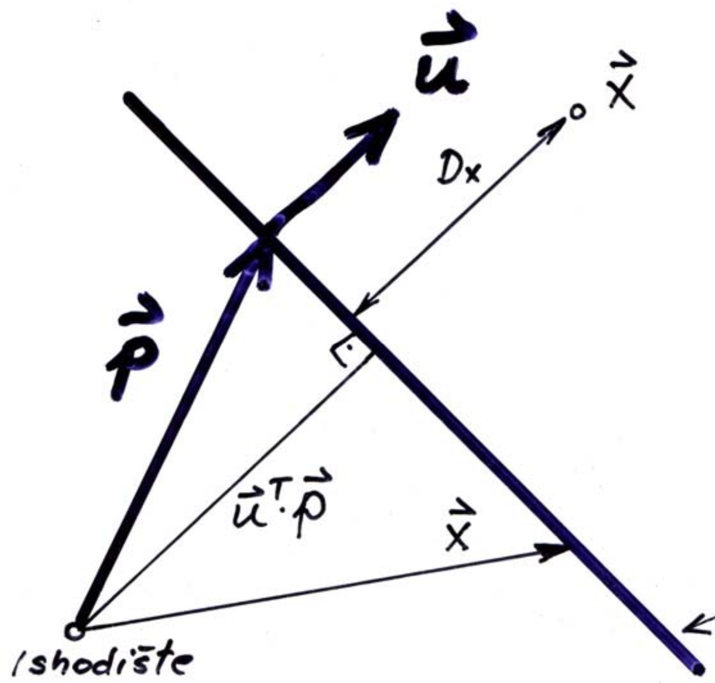
nedefinirano

$d(\vec{x})$ zapišimo u vektorskom obliku:

$$d(\vec{x}) = \vec{W}_0^T \vec{x} + W_{n+1} = 0$$

Geometrijska interpretacija funkcije odlučivanja:





$\vec{u}$ - jedinični vektor normalan na hiperravninu u točki $\vec{p}$

hiperravnina

$$\vec{W}_0^T \cdot \vec{X} + W_{n+1} = 0$$

Jednadžba hiperravnine:

$$\vec{u}^T(\vec{x} - \vec{p}) = 0$$

$$\vec{W}_0^T \vec{X} + W_{n+1} = 0 \quad / \|\vec{W}_0\|$$

$$\|\vec{W}_0\| = \sqrt{W_1^2 + W_2^2 + \ldots + W_n^2}$$

$$\frac{\vec{W}_0^T \vec{X}}{\|\vec{W}_0\|} = -\frac{W_{n+1}}{\|\vec{W}_0\|}$$

usporedimo!

$$\vec{u}^T \vec{x} = \vec{u}^T \cdot \vec{p}$$

$$\vec{u} = \frac{\vec{W}_0}{\|\vec{W}_0\|} \quad i \quad \vec{u}^T \cdot \vec{p} = -\frac{W_{n+1}}{\|\vec{W}_0\|}$$

$$\vec{u} = \frac{\vec{W_0}}{\|\vec{W_0}\|}$$

→ pokazuje orijentaciju hiperravnine;
ako je neka komponenta od $\vec{u}$ jednaka 0 onda je hiperravnina paralelna s odgovarajućom koordinatnom osi

Apsolutna vrijednost $\vec{u}^T \cdot \vec{p}$:

$|\vec{u}^T \cdot \vec{p}|$ **predstavlja** udaljenost hiperravnine od ishodišta.

$$D_u = \frac{|W_{n+1}|}{\|\vec{W_0}\|}$$

Posljedice:

- Budući da je $\vec{u} = \frac{\vec{W_0}}{\|\vec{W_0}\|}$ ispitivanjem vektora težinskih koeficijenata $\vec{W_0}$ moguće je utvrditi da li je hiperravnina paralelna s bilo kojom koordinatnom osi.
- Ako je $W_{n+1} = 0$ onda hiperravnina prolazi kroz ishodište koordinatnog sustava
- Udaljenost točke $\vec{X}$ od hiperravnine je $D_X = |\vec{u}^T \cdot \vec{p} - \vec{u}^T \cdot \vec{X}|$

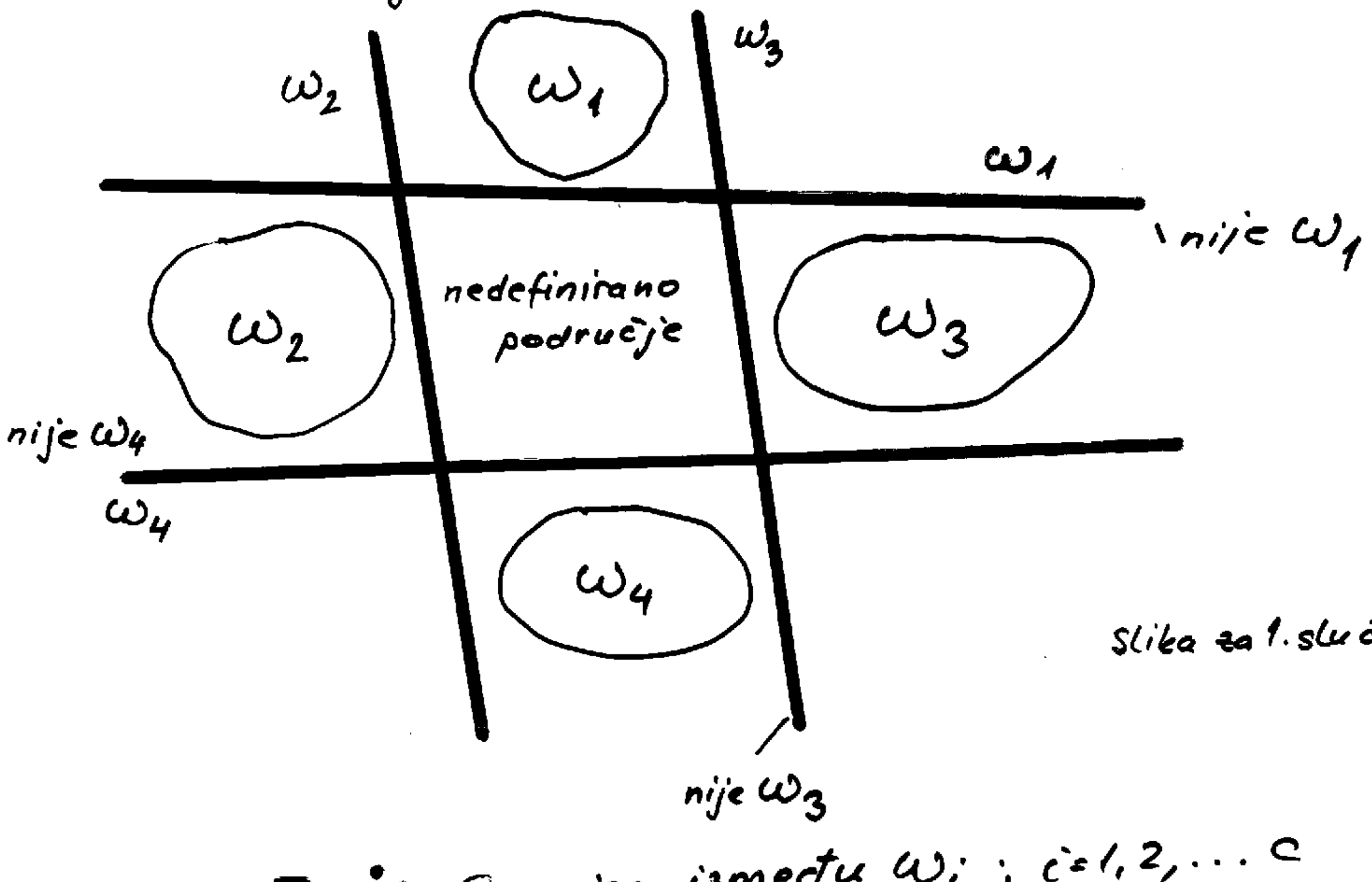
$$= \left| \frac{\vec{W_0}^T \cdot \vec{p}}{\|\vec{W_0}\|} - \frac{W_{n+1}\vec{X}}{\|\vec{W_0}\|} \right|$$

SLUČAJ: VIŠE RAZREDA
$M > 2$

- više pristupa rješavanju problema linearnog klasifikatora za $M > 2$

$M = c \qquad c > 2$

Primjer: problem se može reducirati na c problema klasifikacije u dva razreda u kojem se i-ti problem rješava linearnom funkcijom odlučivanja koja odvaja uzorke razreda ω_i od svih ostalih koji ne pripadaju razredu ω_i.



Slika za 1. slučaj

1. Slučaj: Granica između ω_i; $i = 1, 2, \dots c$ i preostalih razreda

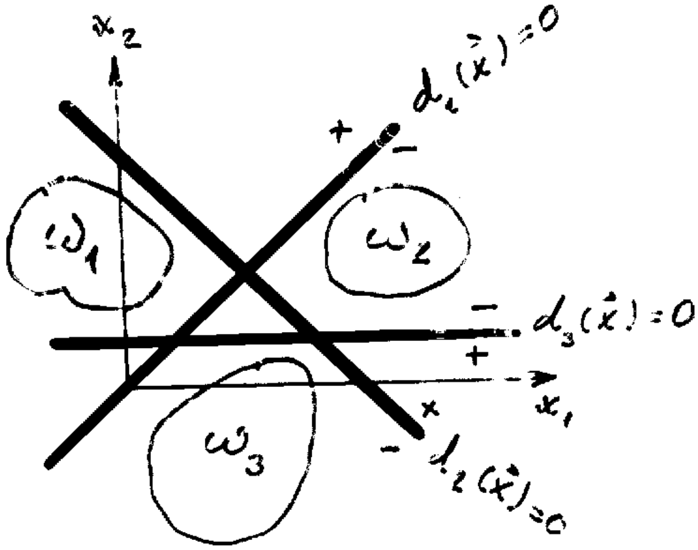
$$d_i(\vec{x}) = W_{i1}x_1 + W_{i2}x_2 + \dots + W_{in}x_n + W_{i,n+1} = 0$$

$$d_i(\vec{x}) = \vec{W_i}^T \vec{X} + W_{i\,n+1}$$

1. slučaj:

Svaki je razred uzoraka separabilan od ostalih razreda jednom decizijskom ravninom

$$d_i(\vec{x}) = \vec{w}_i^T \vec{x} = \begin{cases} >0 & \text{if } \vec{x} \in \omega_i \\ <0 & \text{if } \vec{x} \notin \omega_i \end{cases}$$

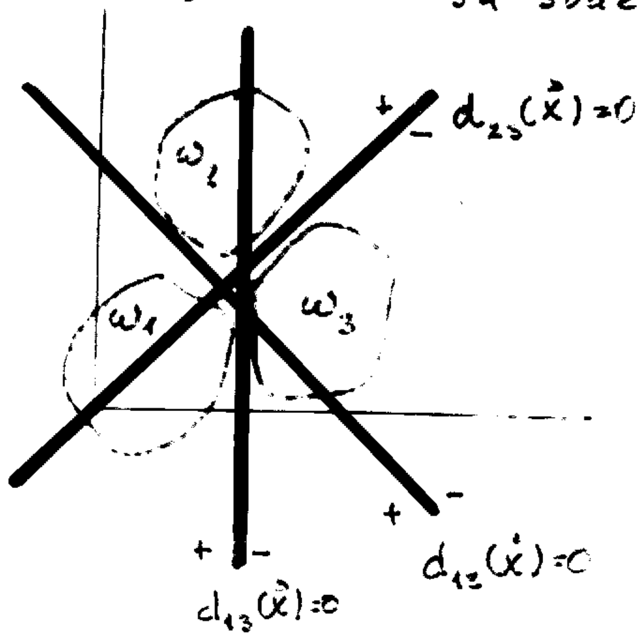


2. slučaj:

Svaki razred uzoraka je separabilan sa svakim pojedinim (drugim) razreda i to jednom decizijskom ravninom.

/razredi su po parovima separabilni:

$M(M-1)/2$ decizijskih ravnina



3. slučaj: Postoji M decizijskih funkcija

$$d_k(\vec{x}) = \vec{w}_k^T \vec{x}, \quad k = 1, 2, \ldots, M \quad \text{sa svojstvom}$$

da $\vec{x}$ pripada razredu ω_i ako

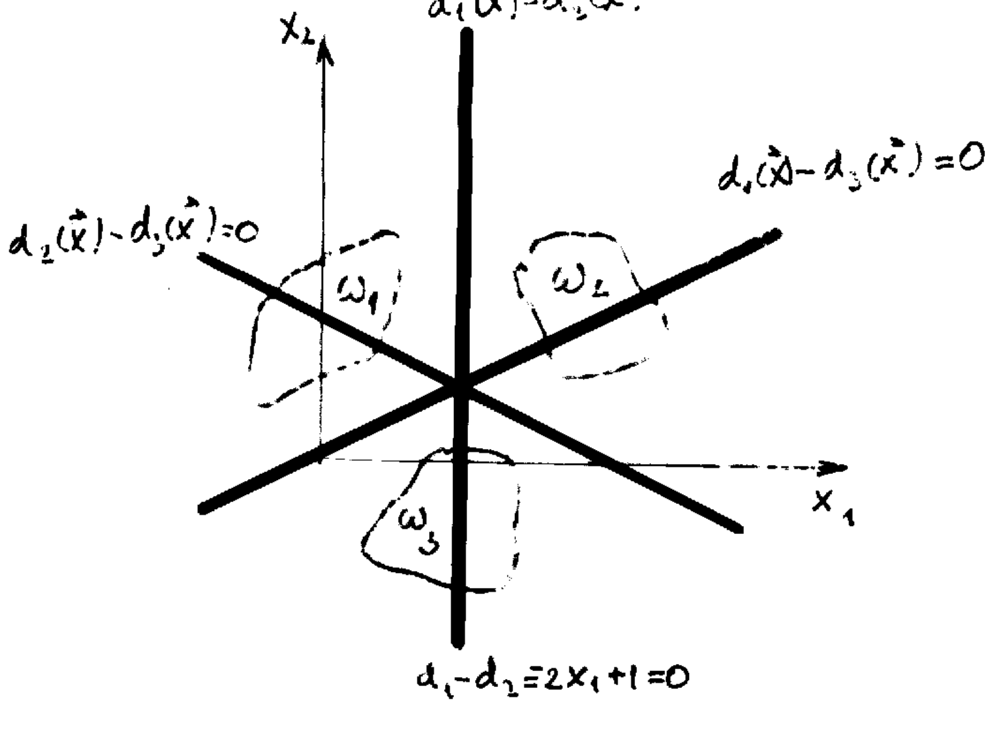
$$d_i(\vec{x}) > d_j(\vec{x}) \quad \text{za sve } j \neq i$$

To je poseban slučaj **2. slučaja** zato što možemo definirati sljedeće:

$$d_{ij}(\vec{x}) = d_i(\vec{x}) - d_j(\vec{x})$$
$$= (\vec{w}_i - \vec{w}_j)^T \vec{x}$$
$$= \vec{w}_{ij}^T \vec{x} \quad ; \quad \vec{w}_{ij} = \vec{w}_i - \vec{w}_j$$

ako je $d_i(\vec{x}) > d_j(\vec{x})$ za sve $j \neq i$, tada je $d_{ij}(\vec{x}) > 0$ za sve $j \neq i$, što znači:

ako su razredi separabilni za 3. slučaj onda su automatski separabilni i za 2. slučaj.



2. slučaj :

Upotrijebi $\frac{c(c-1)}{2}$ linearnih funkcija odlučivanja tako da sa svakom funkcijom odvojiš par razreda
$\overset{}{\underset{2}{\|}}$

Granica između ω_i i ω_j je zadana s :

$$d_{ij}(\vec{x}) = W_{ij1} X_1 + W_{ij2} X_2 + \cdots + W_{ijn} X_n + W_{ijn+1} = 0$$

3. Slučaj :

$$d_{ij}(\vec{x}) = d_i(\vec{x}) - d_j(\vec{x}) =$$

$$= (W_{i1} - W_{j1}) X_1 + (W_{i2} - W_{j2}) X_2 + \ldots +$$

$$(W_{in} - W_{jn}) \cdot X_n + (W_{in+1} - W_{jn+1}) = 0$$

$$d_i(\vec{X}) = \vec{W}_i^T \vec{X} + W_{in+1}$$

ako je $d_i(\vec{x}) > d_j(\vec{x})$; $j = 1, 2, \ldots ; j \neq i$

onda
$\vec{X} \in \omega_i$

klasifikator $\rightarrow$ Linearni stroj (engl. linear machine)

$$(\vec{W}_i - \vec{W}_j)^T \vec{X} + (W_{in+1} - W_{jn+1}) = 0$$

ODREĐIVANJE FUNKCIJE ODLUČIVANJA → UČENJE ILI VJEŽBANJE

Problem oblikovanja linearnog klasifikatora:
odrediti koeficijente linearne funkcije odlučivanja:

$$\vec{W} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \quad i \quad w_{n+1} \quad ?$$

$$\boxed{d(\vec{x}) = \vec{W}^T \vec{X} + w_{n+1}}$$

"Automatizirati" postupak određivanja koeficijenata linearne funkcije odlučivanja:
iterativni postupak **učenja** koeficijenata linearne funkcije odlučivanja uporabom uzoraka iz skupa za učenje (engl. training set).

N uzoraka: $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N$
razvrstani u dva razreda ω_1 i ω_2
Vektori uzoraka $\vec{x}_i$ $i = 1, 2, \dots N$
su "označeni" vektori tj. oni sa poznatom pripadnosti razredu (ω_1 ili ω_2). **UPOTRIJEBIT ĆEMO IH ZA UČENJE $d(\vec{x})$!**

Povećat ćemo dimenzionalnost vektora $\vec{X}$ za jedan

$$\vec{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ 1 \end{bmatrix}$$

radi elegantnijeg zapisa:

$$\boxed{d(\vec{x}) = \vec{W}^T \vec{x}}$$

Uspijemo li odrediti takav vektor težinskih koeficijenata $\vec{W}$ tako da pomoću funkcije $d(\vec{x})$ pravilno razvrstamo sve uzorke (iz skupa za učenje), kažemo da su ω_1 i ω_2 <u>LINEARNO RAZDVOJIVI</u>.

$M=2$

Uzorak $\vec{x}$ je pravilno razvrstan ako za sve $\vec{x}$ iz ω_1 vrijedi $\vec{W}^T\vec{X} > 0$ i ako za sve $\vec{X}$ iz ω_2 vrijedi $\vec{W}^T\vec{X} < 0$

Jedinstven uvjet: $\vec{W}^T\vec{X} > 0$ ako uzorke iz ω_2 pomnožimo s -1 !

Redefiniran problem:

Tražimo vektor koeficijenata $\vec{w}$ linearne funkcije odlučivanja tako da vrijedi

$$\vec{w}^T \vec{x} > 0 \quad \text{za sve uzorke}$$

$\vec{x}$ iz skupa uzoraka za učenje.

/ PAZI: uzorci $\vec{x} \in \omega_2$ su pomnoženi s -1 /

odnosno

$$[X]\,\vec{w} > \vec{0} \quad \text{za sve uzorke } \vec{x}$$

$$[X] = \begin{bmatrix} \vec{x}_1^T \\ \vec{x}_2^T \\ \vdots \\ \vec{x}_N^T \end{bmatrix}$$

$[X]$ je matrica svih uzoraka iz skupa za učenje, s tim da su uzorci iz ω_2 pomnoženi s -1.

$\vec{w} = (w_1, w_2, \ldots, w_n, w_{n+1})^T$

$\vec{0}$ – nulti vektor

PRIMJER 1:

$$\vec{X}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} ; \quad \vec{X}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \begin{matrix} \vec{X}_1 \in \omega_1 \\ \vec{X}_2 \in \omega_1 \end{matrix}$$

$$\vec{X}_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} ; \quad \vec{X}_4 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \begin{matrix} \vec{X}_3 \in \omega_2 \\ \vec{X}_4 \in \omega_2 \end{matrix}$$

- POVEĆAJMO DIMENZIONALNOST VEKTORA:

$$\vec{X}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} ; \quad \vec{X}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} ; \quad \vec{X}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\vec{X}_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

- POMNOŽIMO SVE UZORKE IZ RAZREDA ω_2 S -1

$$\vec{X}_3 = \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix} ; \quad \vec{X}_4 = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$$

- Oblikujmo matricu $[X]$

$$[X]_{4\times 3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 0 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

-

$$[X]\vec{w} > \vec{0}$$

$$\underset{4\times 3}{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 0 & -1 \\ -1 & -1 & -1 \end{bmatrix}} \cdot \underset{3\times 1}{\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}} > \underset{4\times 1}{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}}$$

vektor $\vec{w}$ koji zadovoljava sustav linearnih nejednadžbi

$$[X]\cdot\vec{w} > \vec{0}$$

nazivamo <u>razdvojni vektor</u>.

A) GRADIJENTNI POSTUPCI ODREĐIVANJA RAZDVOJNOG VEKTORA

$d(\vec{x})$ - funkcija vektorskog argumenta

$\quad$ vektor!

Općenito: $f(\vec{y})$

$$\vec{y} = (y_1, y_2, \ldots, y_n)^T$$

Gradijent funkcije vektorskog argumenta:

$$\text{grad } f(\vec{y}) = \frac{d f(\vec{y})}{d\vec{y}} = \begin{bmatrix} \frac{df}{dy_1} \\ \frac{df}{dy_2} \\ \vdots \\ \frac{df}{dy_n} \end{bmatrix}$$

- GRADIJENT SKALARNE FUNKCIJE VEKTORSKOG ARGUMENTA JE **VEKTOR**

- svaka komponenta gradijenta predstavlja veličinu promjene funkcije u smjeru komponente vektora

VRIJEDI :

- **Povećanje argumenta u smjeru pozitivnog gradijenta funkcije f dovodi nas do <u>maksimuma funkcije f</u>**

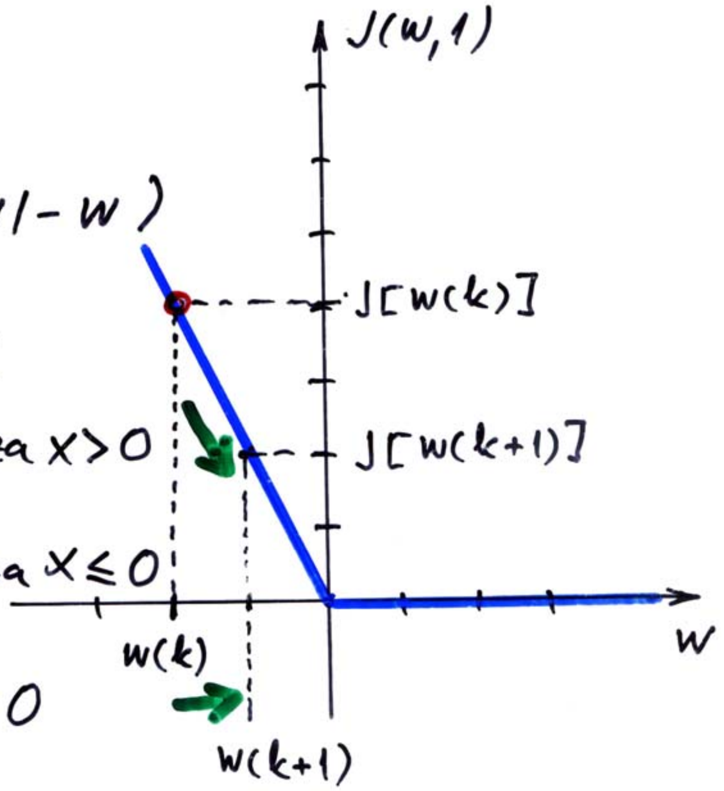
- **povećanje argumenta u smjeru negativnog gradijenta dovodi nas do <u>minimuma funkcije f</u>**

<u>PRIMJER **2**:</u>

Funkcija $J(w,1) = (|w| - w)$

$$\frac{\partial J(w,1)}{\partial w} = \operatorname{sgn}(w) - 1$$

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{za } x > 0 \\ -1 & \text{za } x \leq 0 \end{cases}$$



$$\frac{\partial J}{\partial w} = -2 \quad \text{ako je } w \leq 0$$

$$= 0 \quad \text{ako je } w > 0$$

$w(k+1) > w(k)$

povećanje argumenta u smjeru negativne vrijednosti ~~komp.~~ gradijenta $\Rightarrow$ prema min.