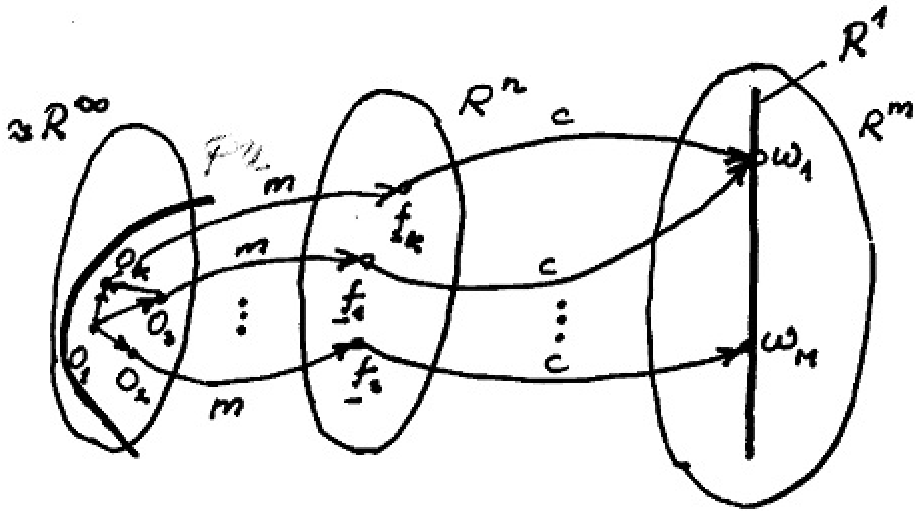
1 Model sustava za raspoznavanje uzoraka

1.1 Formalni model



Slika 1: Formalni model sustava RU

Okolina O : $o_k \in O; v_j \in VO$

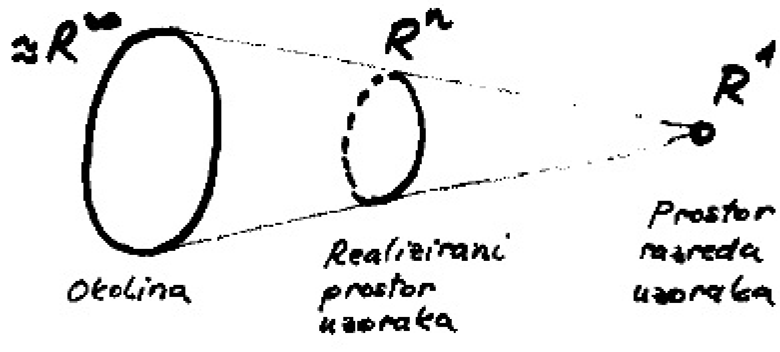
Realizirani prostor uzoraka PU : f_k-uzorci

Prostor razreda uzoraka Ω : $\Omega = \{\omega_1, \omega_2, \ldots, \omega_M\}$

mjerenje m: $m : o_k -> f_k$

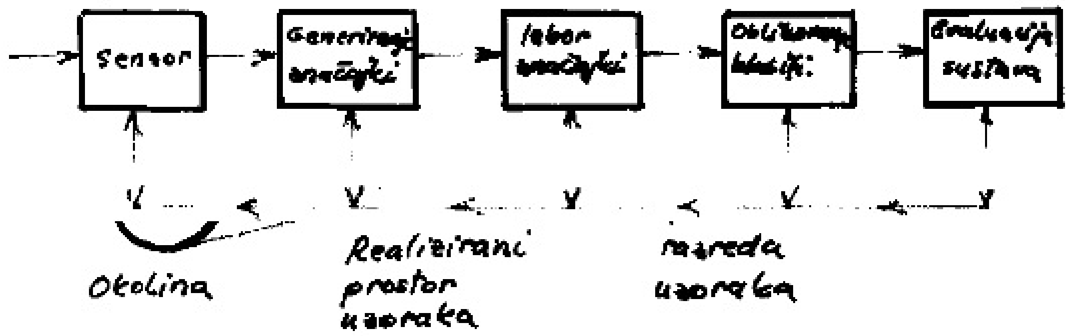
klasifikacija c: $c : f_k -> \omega_i$

Uhrov stožac raspoznavanja - redukcija informacija



Slika 2: Uhrov stožac

1.2 Značajke, vektor značajki i klasifikator



Slika 3: Osnovne faze u postupku oblikovanja sustava RU

značajke - promatramo ih kao slučajne varijable: x_i

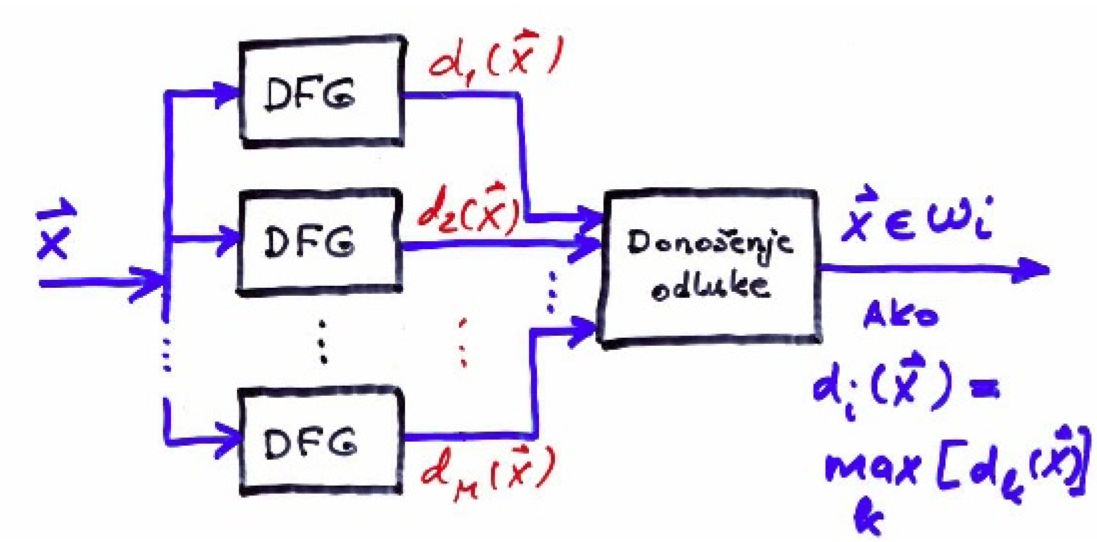
vektor značajki - slučajni vektori $\vec{X}$, $\vec{X} = [x_1, x_2, \ldots, x_n,]^T$

-značajke koje predstavljaju razlike između razreda uzoraka nazivaju se INTERSET značajke
-INTRASET značajke su zajedničke SVIM razredima iz PU (područja uporabe) i ne nose diskriminacijsku informaciju- TAKVE ZNAČAJKE MOGU SE ZANEMARITI
Izbro 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 informacija
- 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 procedure
- Decizijske granice određene su decizijskim funkcijama: $d_1(\vec{X}), d_2(\vec{X}), \ldots, d_M(\vec{X})$, /VAŽNO: d_i je funkcija koja ima za argument VEKTOR a vraća SKALAR/

- Pravilo razvrstavanja:
  Ako $d_i(\vec{X}) > d_j(\vec{X})$ za i, j=1,2,...,M te je $j \neq i$ tada nepoznati uzorak $\vec{X}$ pripada razredu ω_i

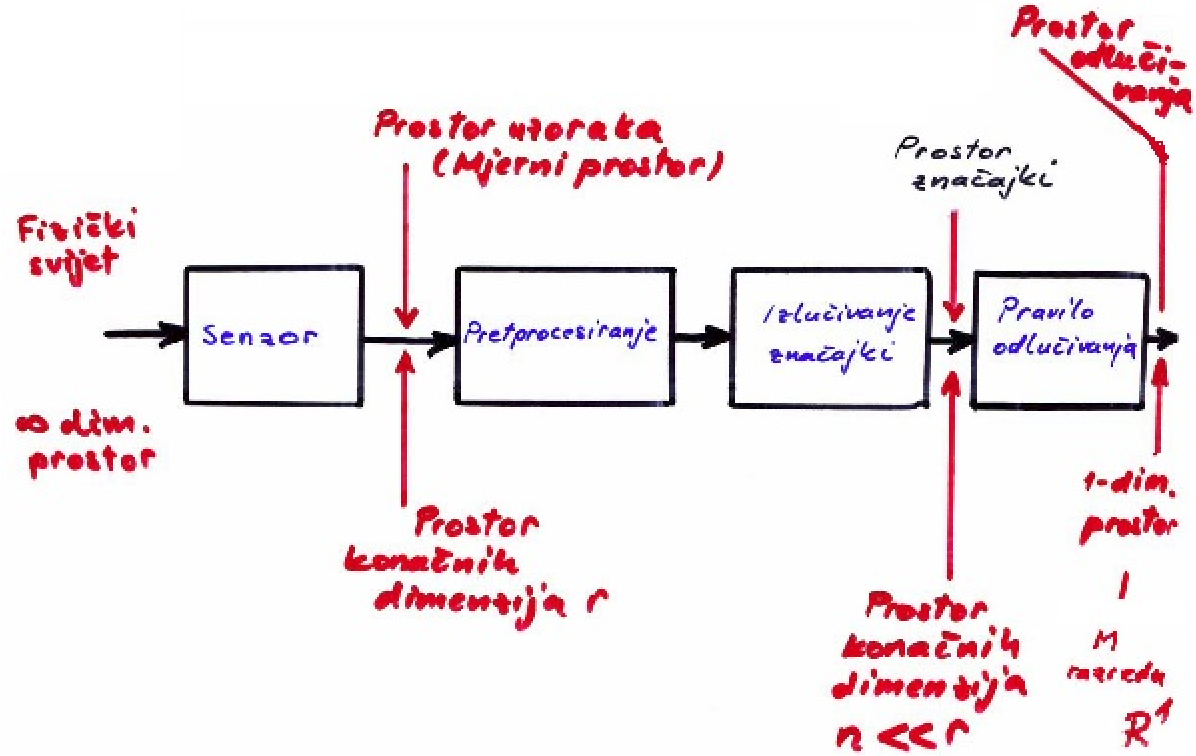
Blok-dijagram klasifikatora



Slika 4: Blok-dijagram klasifikatora

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

Model sustava za raspoznavanje



Slika 5: Model sustava za raspoznavanje

- Vanjski, fizički ("analogni") svijet sadrži praktički ∞ mnogo značajki
  - senzor ili pretvarač pretvara analogni svijet u zapis koji sadrži r (brojčanih) vrijednosti
- Pretprocesiranje : izlučivanje šuma, poboljšanje mjernog podatka
- Prostor značajki: $n << r$, $\vec{X} = (x_1, x_2, \ldots, x_n)^T$
- Prostor odlučivanja - jednodimenzionalni prostor R^1

PRIMJER:

$OB = \{o_k; k = 1, 2, \ldots\}$
o_i - signal EKG-a pacijenta
PU - automatska dijagnoza signala EKG-a u testu opterećenja
Uzorak: $f_k(\vec{X}) -> \vec{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$
Razred objekata:

Normalni EKG

Granični EKG

Nenormalni EKG

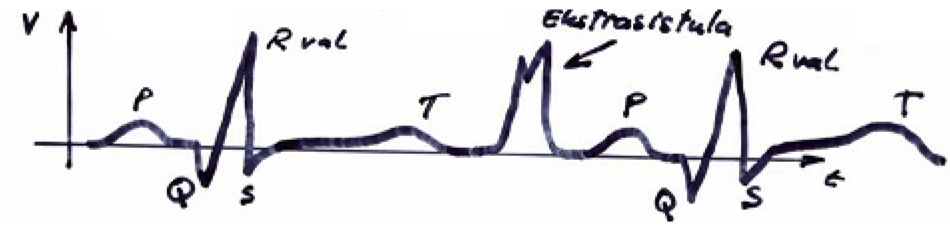
Razred uzoraka:

C_1/ω_1 - Nenormalni EKG

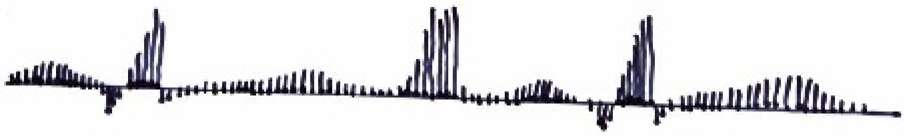
C_2/ω_2 - Granični EKG

C_3/ω_3 - Normalni EKG

Skup uzoraka za učenje ili vježbanje $\mathcal{U}_M = (S_N, \Omega)$
Tip uzorka - jednostavan uzorak
Sustav za automatsku dijagnozu EKG-a /Lesterova dijagnostička metoda/



Slika 6: Vanjski svijet



Slika 7: Uzorkovanje (f_u=800Hz)

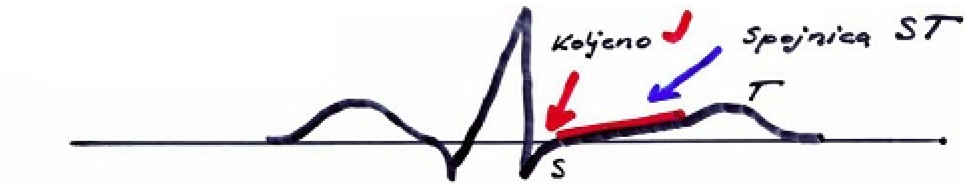
-Pretprocesiranje:

- Elimnacija elektrosistule
- 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

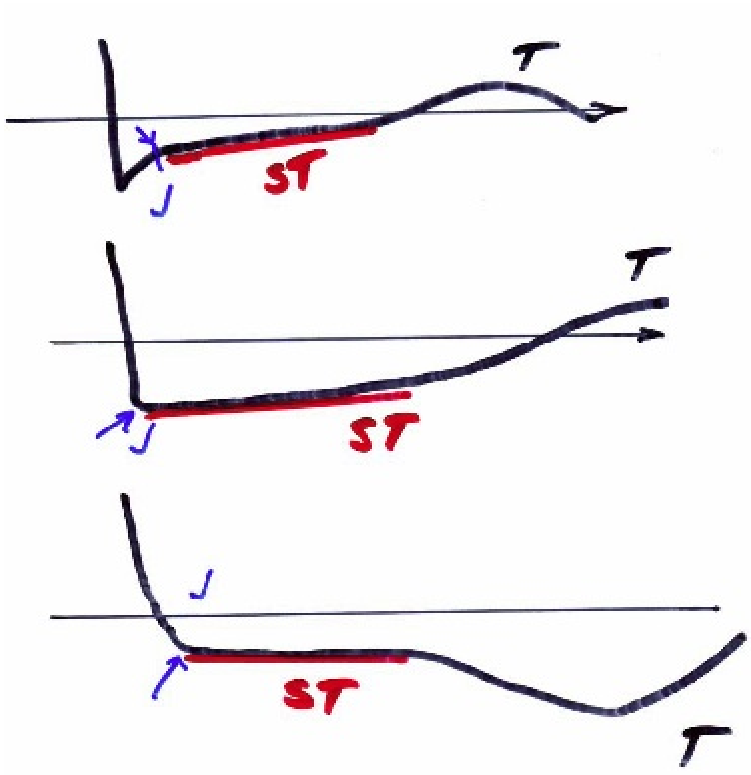
Problem:

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

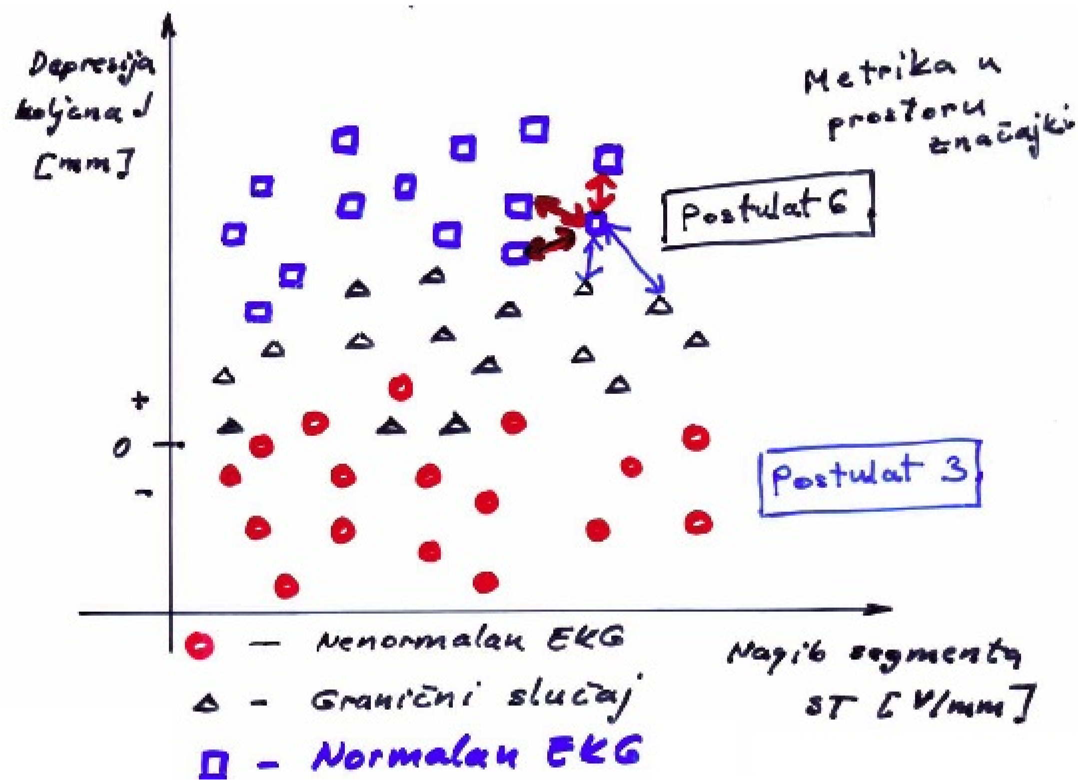


Slika 8: Prikaz koljena

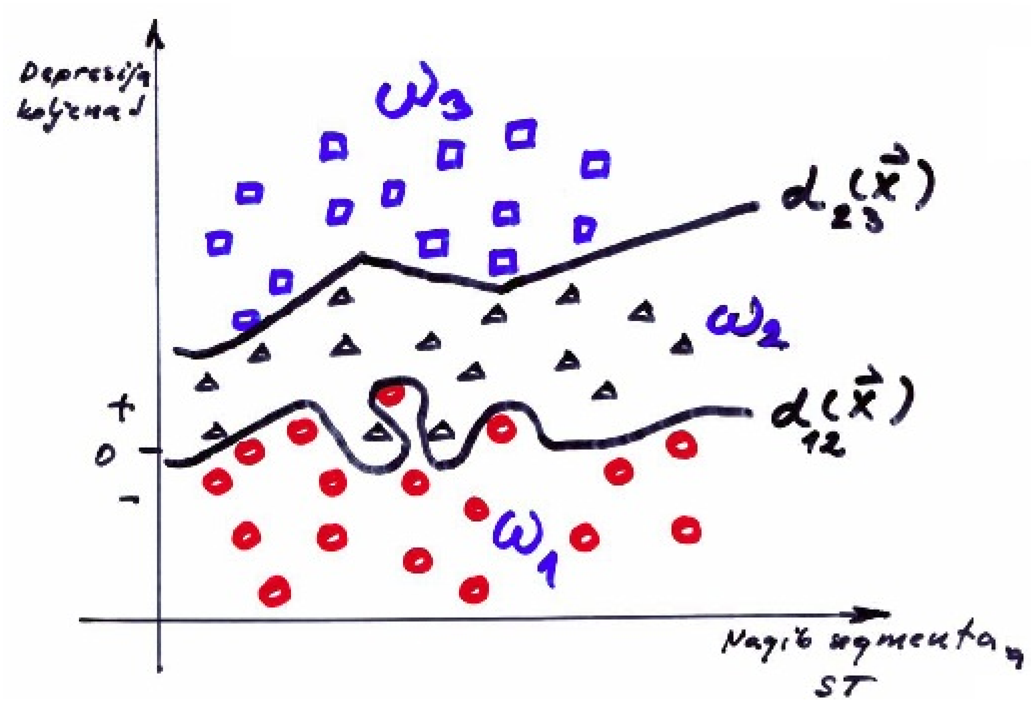
Označeni uzorci: u skladu s postulatom 1
Ω=$\{NenormalanEKG\omega_1$, Granični slučaj ω_2, Normalan EKG ω_3
$= \{\omega_1$, ω_2, ω_3 $M = 3$
$vektorzna\check{c}ajki$=$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$; x_1 - nagib segmenta ST, x_2 - depresija koljena J



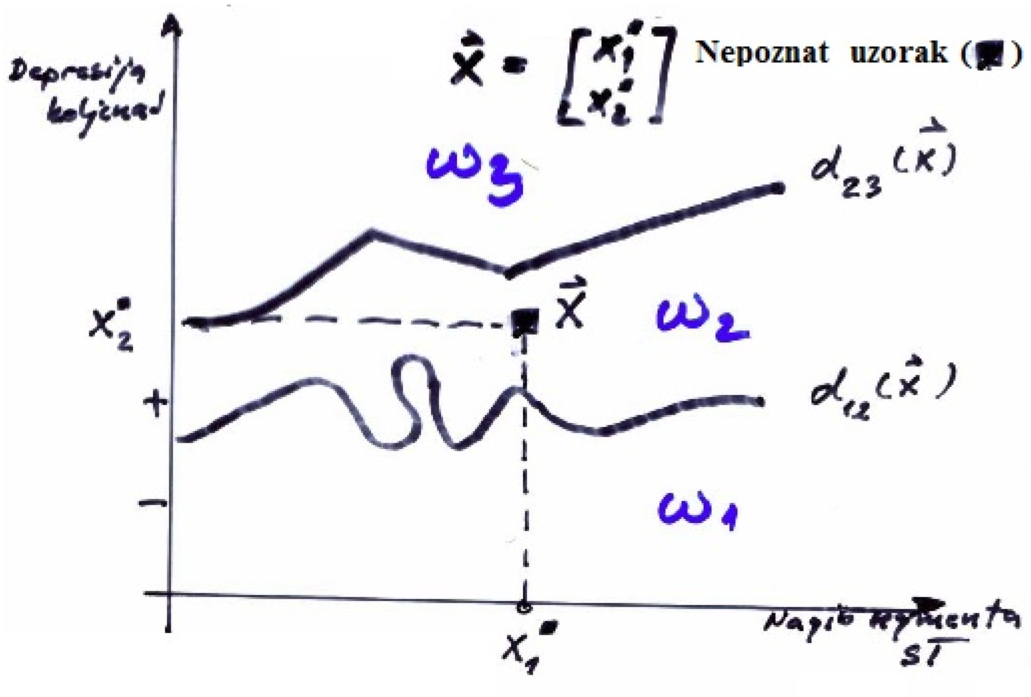
Slika 9: Primjeri koljena J i segmenata ST



Slika 10: Prostor značajki za primjer

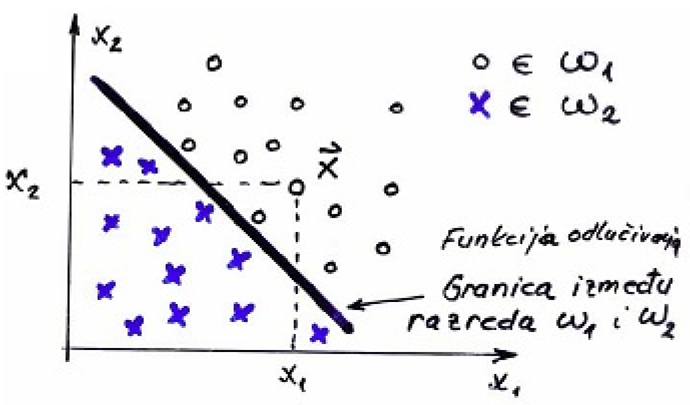


Slika 11: Decizijske funkcije za primjer



Slika 12: Razvrstavanje nepoznatog uzorka

Linearne funkcije odlučivanja
(engl. Linear discriminant functions)



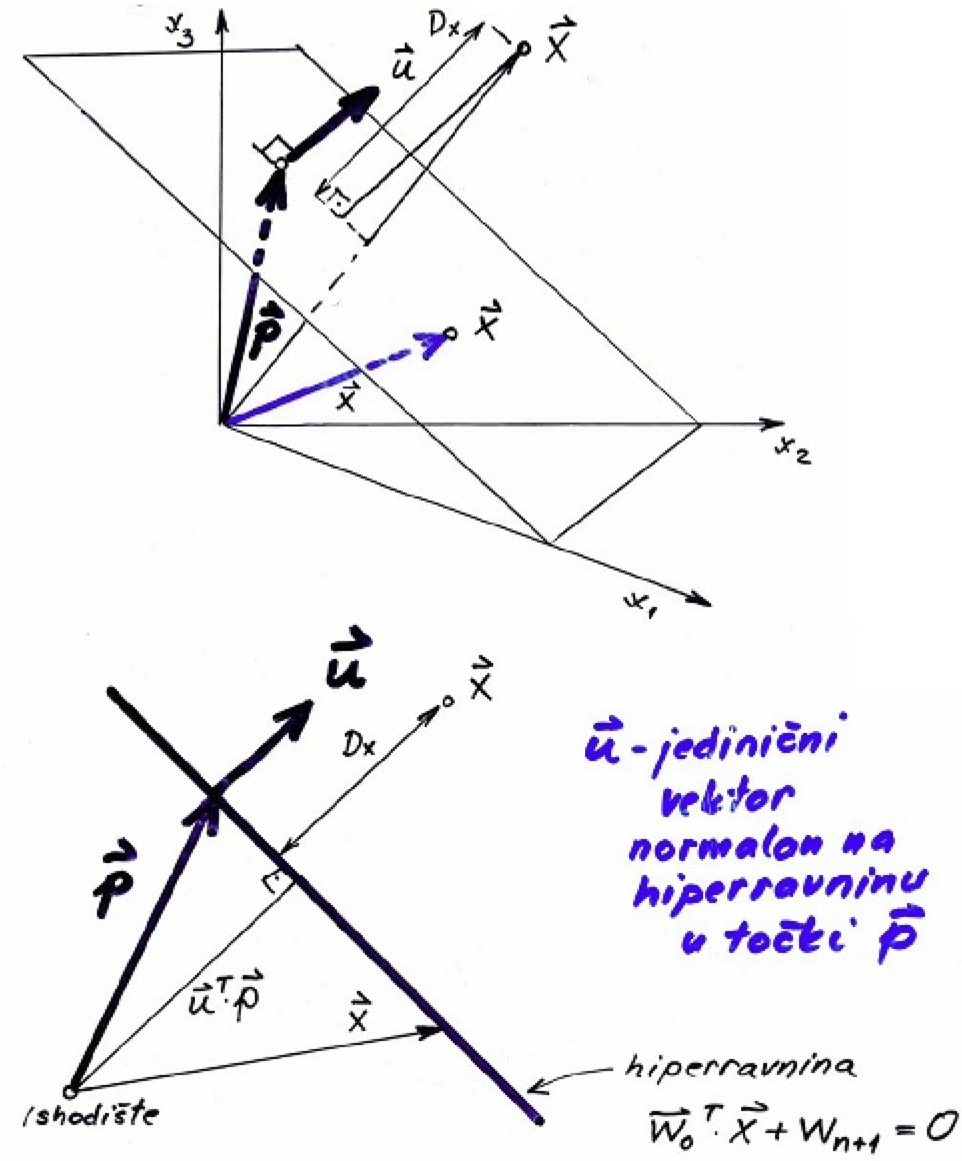
Slika 13: Linearna funkcija odlučivanja

vektor značajki - $\vec{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, M=2

Funkcija odlučivanja kao linearna kombinacija komponenti vektora $\vec{X}$:
$d(\vec{X}) = w_1x_1 + w_2x_2 + \ldots + w_nx_n + w_{n+1}$
za: n=2 funkcija odlučivanaja -> jednadžba pravca
n=3 funkcija odlučivanaja -> jednadžba ravnine
n>3 funkcija odlučivanaja -> hiperravnina
w_i - težinski koeficijenti; i=1,2,...,n
w_{n+1} - pomaknuće, utežnosni prag (engl.bias, threshold weight)

Slučaj 2 razreda (ω_1 i ω_2, M=2)

$d(\vec{X}) = w_1x_1 + w_2x_2 + \ldots + w_nx_n + w_{n+1} = 0$
M=2, ω_1, 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 $d(\vec{X}) = 0$ onda nedefinirano
$d(\vec{X})$ zapišimo u vektorskom obliku: $d(\vec{X}) = \vec{w_0}^T\vec{X} + w_{n+1} = 0$



Slika 14: Geometrijska interpretacija funkcije odlučivanja

Jednadžba hiperravnine:
$\vec{u}^T \cdot (\vec{X} - \vec{p}) = 0$

$\vec{w_0}^T \vec{X} + w_{n+1} = 0 / \|\vec{w_0}\|$, gdje je $\|\vec{w_0}\| = \sqrt{w_1^2 + w_2^2 + \ldots + w_n^2}$

$\frac{\vec{w_0}^T \vec{X}}{\|\vec{w_0}\|} = -\frac{w_{n+1}}{\|\vec{w_0}\|}$

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

$\vec{u} = \frac{\vec{w_0}}{\|\vec{w_0}\|}$ i $\vec{u}^T \cdot \vec{p} = -\frac{w_{n+1}}{\|\vec{w_0}\|}$

$\vec{u} = \frac{\vec{w_0}}{\|\vec{w_0}\|}$ -> pokazuje orijentaciju hiperravnine; ako je neka komponenta od $\vec{u}$ jednaka 0 onda je hiperravnina paralelna s odgovarajućom koordinatnom osi

Apsolutna vrijednost $\vec{u}^T \cdot \vec{p}$:
$\left|\vec{u}^T \cdot \vec{p}\right|$ predstavlja udaljenost hiperravnine od ishodišta.
$D_u = \frac{|w_{n+1}|}{\|\vec{w_0}\|}$
Posljedice:

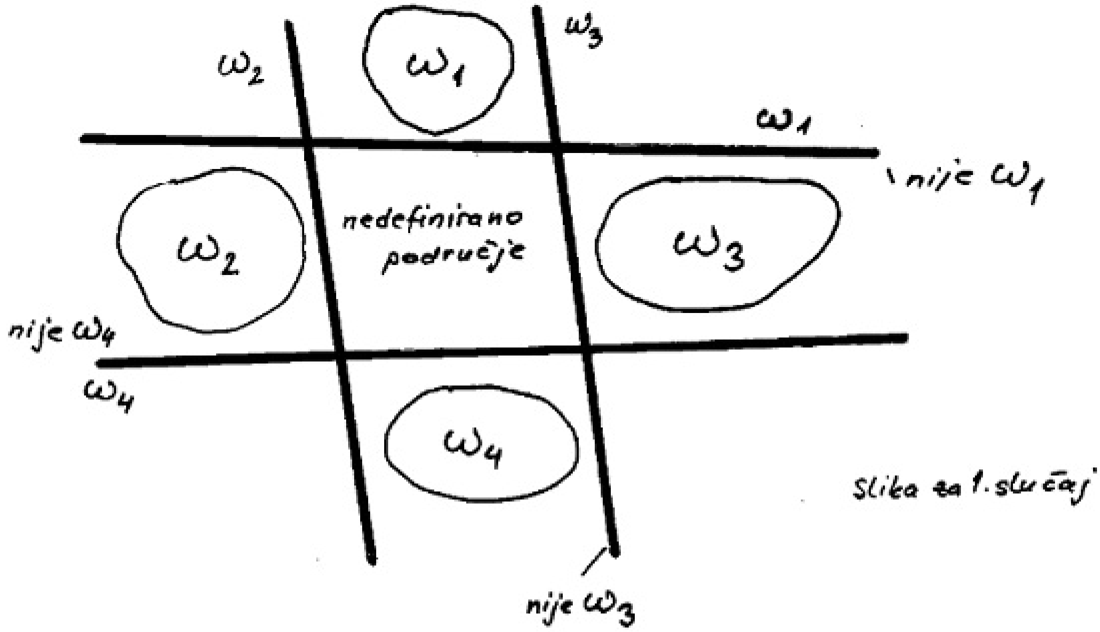
- Budući da je $\vec{u} = \frac{\vec{w_0}}{\|\vec{w_0}\|}$ ispitivanjem vektora težinskih koeficijenata $\vec{w_0}$ moguće je utvrditi da li je hiperravnina paralelna s bilo kojom koordinatnom osi
- ako je $w_{n+1} = 0$ onda hiperravnina prolazi kroz ishodište koordinatnog sustava
- Udaljenost točke $\vec{X}$ od hiperravnine je $D_x = \left|\vec{u}^T \cdot \vec{p} - \vec{u}^T \cdot \vec{X}\right| = \left|\frac{\vec{w_0}^T \cdot \vec{p}}{\|\vec{w_0}\|} - \frac{w_{n+1}\vec{X}}{\|\vec{w_0}\|}\right|$

SLUČAJ: VIŠE RAZREDA M>2

- više pristupa rješavanju problema linearnog klasifikatora za M>2

M=c, c>2

Primjer problem se može reducirati na c problema klasifikacije u dva razreda u kojem se i-ti problem rješava linearnom funkcjiom odlučivanja koja odvaja uzorke razreda ω_i od svih ostalih razreda ω_i.



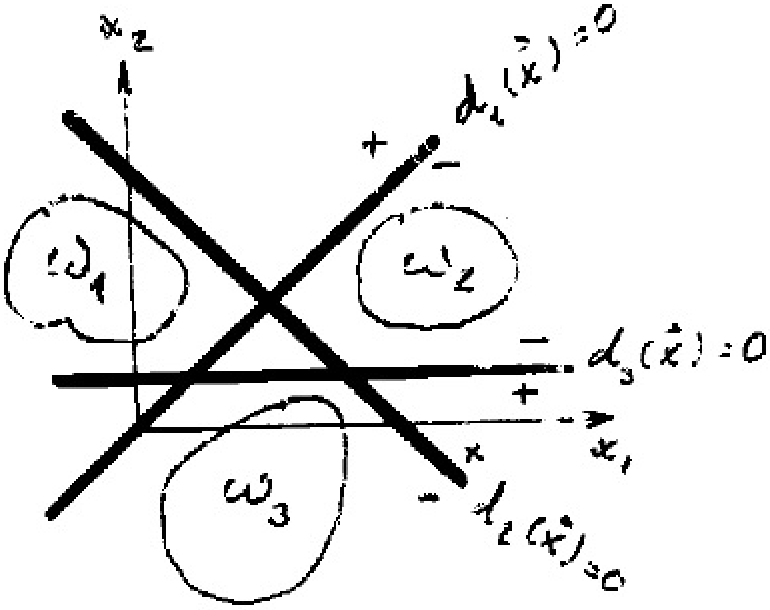
Slika 15: Klasifikacija više razreda

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

$d_i(\vec{X}) = w_{i1}x_1 + w_{i2}x_2 + \ldots + w_{in}x_n + w_{in+1} = 0$

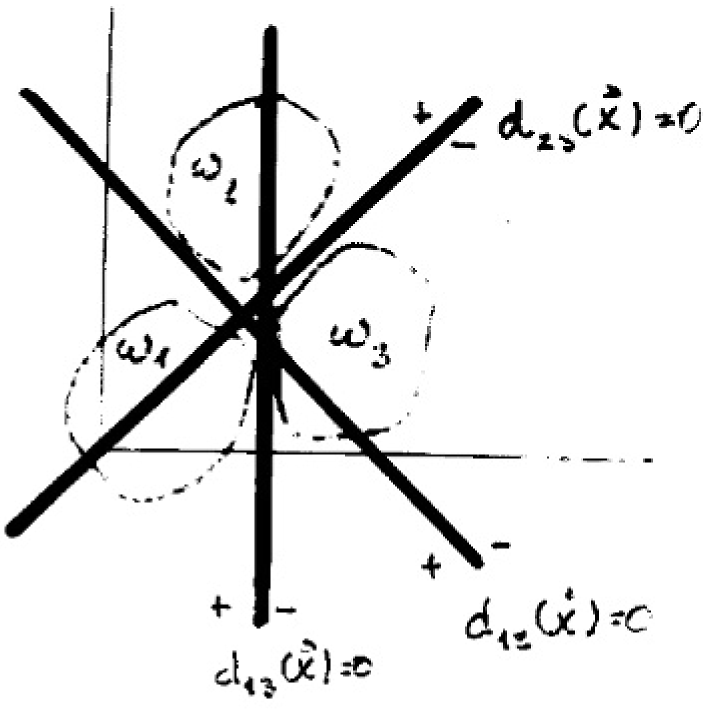
$d_i(\vec{X}) = \vec{w_i}^T \vec{X} + w_{in+1}$

Svaki je razred uzoraka separabilan od ostalih razreda jednom decizijskom ravninom $d_i(\vec{X}) = \vec{w_i}^T \vec{X} = \begin{cases} > 0 \text{ za } \vec{X} \in \omega_i \\ < 0 \text{ za } \vec{X} \notin \omega_i \end{cases}$



Slika 16: 1. slučaj

2. slučaj: Svaki razred uzoraka je separabilan sa svakim pojedinim (drugim razredom) i to jednom decizijskom ravninom /razredi su po parovima separabilni: M(M-1)/2 decizijskih ravnina/

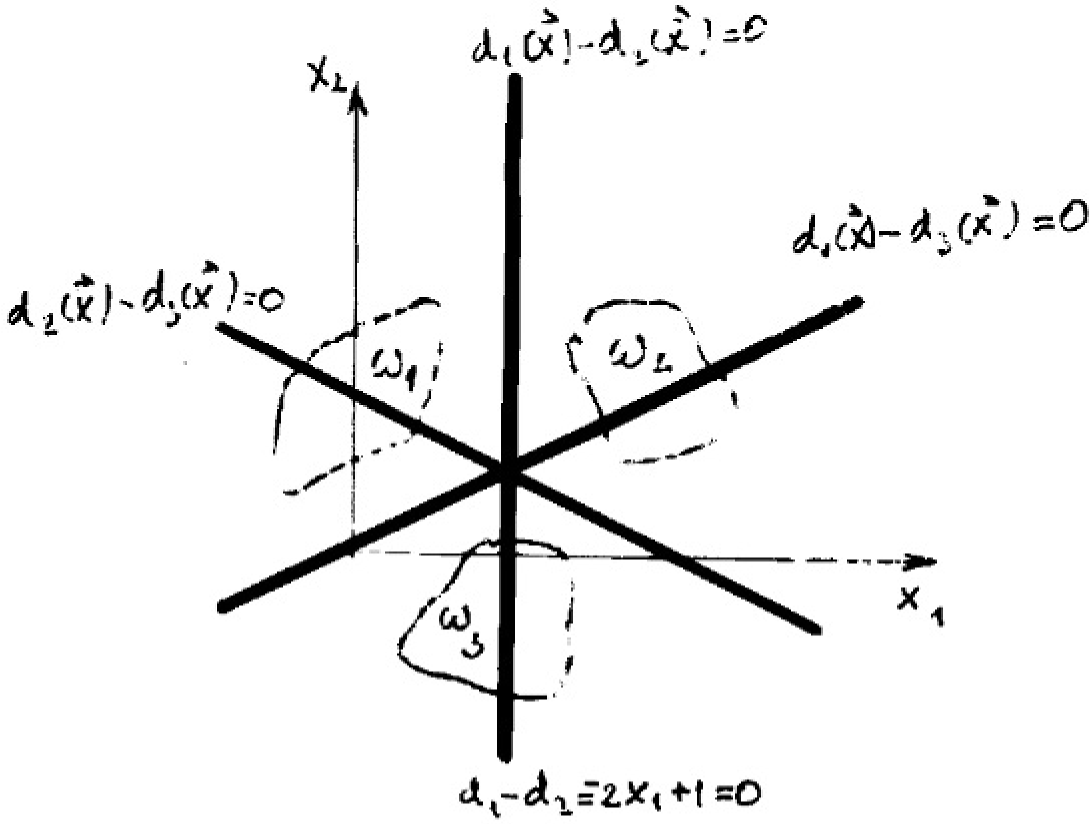


Slika 17: 2. slučaj

Upotrijebi $\frac{c(c-1)}{2}$ linearnih funkcija odlučivanja tako da svakom funkcijom odvojiš par (2) razreda
Granica između ω_i i ω_j je zadana s: $d_ij(\vec{X}) = w_{ij1}x_1 + w_{ij2}x_2 + \ldots + w_{ijn}x_n + w_{ij_{n+1}} = 0$

3. slučaj: Postoji M decizijskih funkcija $d_k(\vec{X}) = \vec{w_k}^T\vec{X}$, $k = 1, 2, \ldots, M$ sa svojstvom da $\vec{X}$ pripada razredu ω_i ako $d_i(\vec{X}) > d_j(\vec{X})$ za sve $j \neq i$
To je poseban slučaj 2. slučaja zato što možemo definirati $d_{ij}(\vec{X}) = d_i(\vec{X}) - d_j(\vec{X}) = (\vec{w_i} - \vec{w_j})^T\vec{X} = \vec{w_{ij}}^T\vec{X}$; $\vec{w_{ij}} = \vec{w_i} - \vec{w_j}$ 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.

$d_{ij}(\vec{X}) = d_i(\vec{X}) - d_j(\vec{X}) = (w_{i1} - w_{j1})x_1 + (w_{i2} - w_{j2})x_2 + \ldots + (w_{in} - w_{jn})x_n + (w_{in+1} - w_{jn+1}) = 0$
$d_i(\vec{X}) = \vec{w_i}^T \cdot \vec{X} + w_{in+1}$ ako je $d_i(\vec{X}) > d_j(\vec{X})$; $i = 1, 2, \ldots; j \neq i$ onda $\vec{X} \in \omega_i$; klasifikator -> linearni sroj (engl. linear machine)
$(\vec{w_i} - \vec{w_j})^T\vec{X} + (w_{in+1} - w_{jn+1}) = 0$



Slika 18: 3. slučaj

ODREĐIVANJE FUNKCIJE ODLUČIVANJA -> UČENJE ILI VJEŽBANJE

Problem oblikovanja lineranog klasifikatora: **odrediti koeficijente linearne funkcije odlučivanja**:

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

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

"Automatizirati" postupak određivanja koeficijenata linearne funkcije odlučivanja:

iterativni postupak **učenja** koeficijenata linerane funkcije odlučivanja uporabom uzoraka iz skupa za učenje(engl. training set).

N uzoraka: $\vec{X_1}, \vec{X_2}, \ldots, \vec{X_N}$ razvrstani u dva razreda ω_i i ω_2

Vektori uzoraka $\vec{X_i}$, $i = 1, 2, \ldots, N$ su "označeni" vektori, tj. oni sa poznatom pripadnosti razredu (ω_i ili ω_2). **UPOTRIJEBIT ĆEMO IH ZA UČENJE** $d(\vec{X})$**!**

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

$$\vec{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ 1 \end{bmatrix} \text{ radi elegantnijeg zapisa: } d(\vec{X}) = \vec{W}^T\vec{X}$$

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 LINEARNO RAZDVOJIVI.

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 pmnožimo s -1!

Redefiniran problem: Tražimo vektor koeficijenata $\vec{W}$linearne funkcije odlučivanja tako da vrijedi $\vec{W}^T\vec{X} > 0$ za sve uzorke za učenje. /PAZI: uzorci $\vec{X} \in \omega_2$ su pomnoženi s -1/

odnosno $[X]\vec{W}>$ 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.

$\vec{W} = (w_1, w_2, \ldots, w_n, w_(n+1)^T$ i $\vec{0}$ - nulti vektor

PRIMJER 1:

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

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

-POVEĆAJMO DIMENZIONALNOST VEKTORA:

$$\vec{X}_1 = \begin{bmatrix}0\\0\\1\end{bmatrix};\ \vec{X}_2 = \begin{bmatrix}0\\1\\1\end{bmatrix};\ \vec{X}_3 = \begin{bmatrix}1\\0\\1\end{bmatrix};\ \vec{X}_4 = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

POMONOŽIMO SVE UZORKE IZ RAZREDA ω_2 S -1:

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

-Oblikujmo matricu [X]

$$[X] = \begin{bmatrix}0 & 0 & 1\\0 & 1 & 1\\-1 & 0 & -1\\-1 & -1 & -1\end{bmatrix}$$

-$[X]\vec{W} > \vec{0}$

$$\begin{bmatrix}0 & 0 & 1\\0 & 1 & 1\\-1 & 0 & -1\\-1 & -1 & -1\end{bmatrix}_{4\times3} \cdot \begin{bmatrix}w_1\\w_2\\w_3\end{bmatrix}_{3\times1} > \begin{bmatrix}0\\0\\0\\0\end{bmatrix}_{4\times1}$$

Vektor $\vec{W}$ koji zadovoljava sustav linearnih nejednadžbi $[X]\vec{W} > \vec{0}$ nazivamo <u>razdvojni vektor</u>.

GRADIJENTNI POSTUPCI ODREĐIVANJA RAZDVOJNOG VEKTORA

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

Općenito: $f(\vec{Y})$, $\vec{Y} = (y_1, y_2, \ldots, y_n)^T$

Gradijent funkcije vektorskog argumenta:

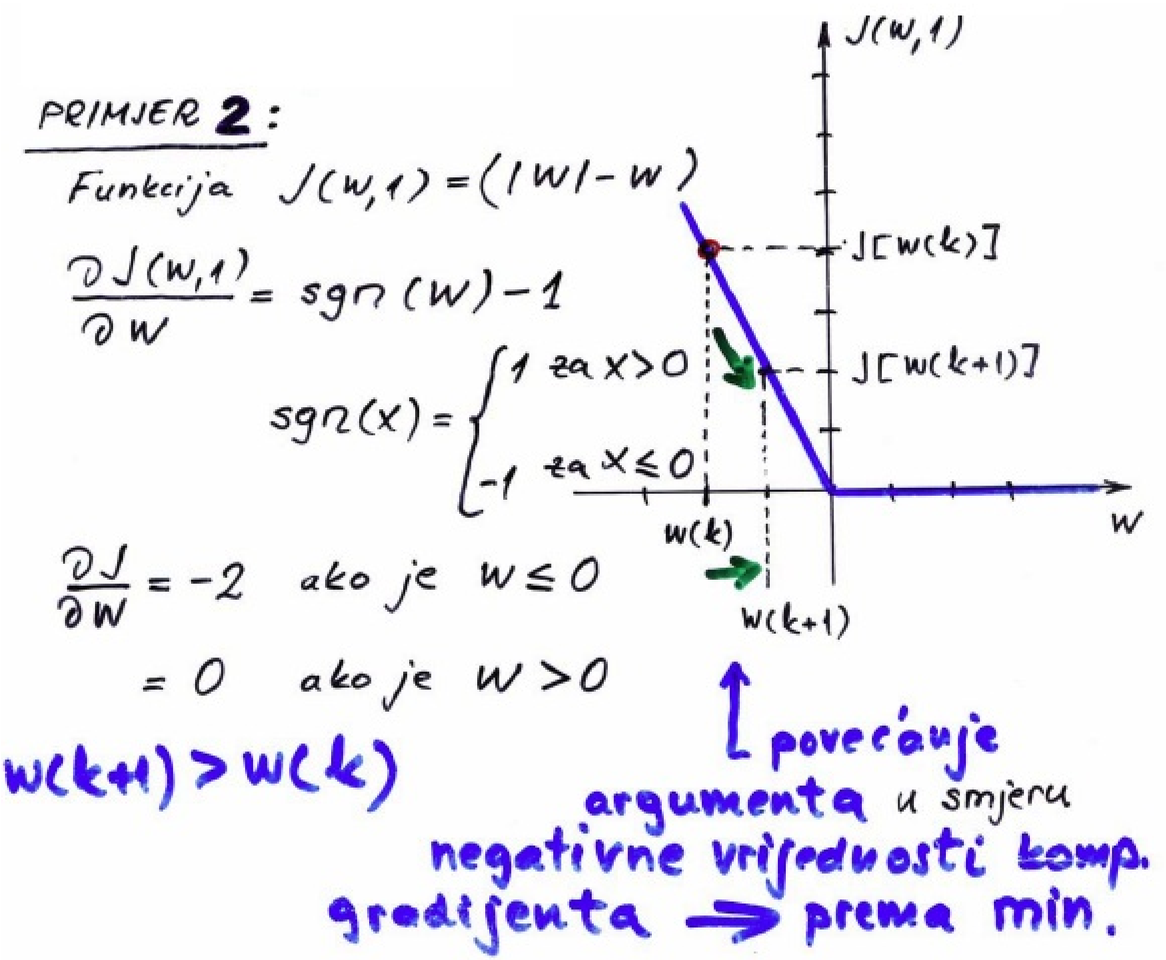
$$\text{grad } f(\vec{Y}) = \frac{df(\vec{Y})}{d\vec{Y}} = \begin{bmatrix}\frac{df}{dy_1}\\ \frac{df}{dy_2}\\ \vdots\\ \frac{df}{dy_n}\end{bmatrix}$$

-GRADIJENT SKALARNE FUNKCIJE VEKTORSKOG ARUGUMENTA JE <u>VEKTOR</u>

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

VRIJEDI:

- Povećanje argumenta u smjeru pozitivnog gradijenta funkcije f dovodi nas do maksimuma funkcije f
- Povećanje argumenta u smjeru negativnog gradijenta funkcije f dovodi nas do minimuma funkcije f



Slika 19: Primjer gradijente metode