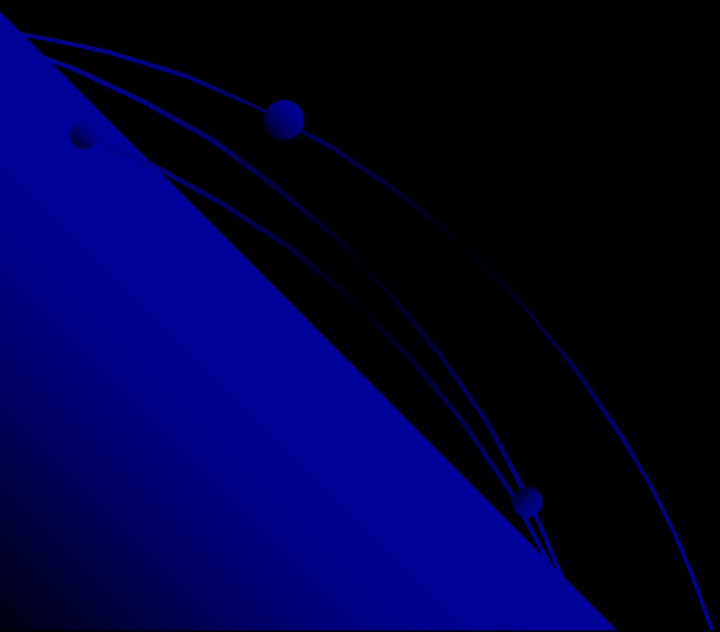


Artificial Neural Networks

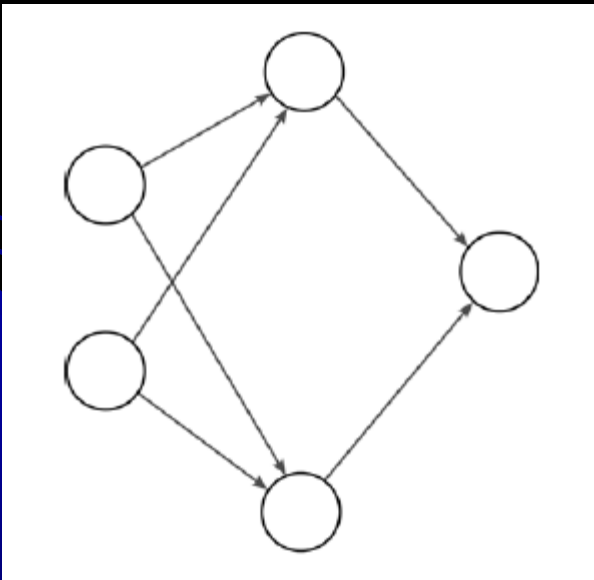
Umelé neurónové siete



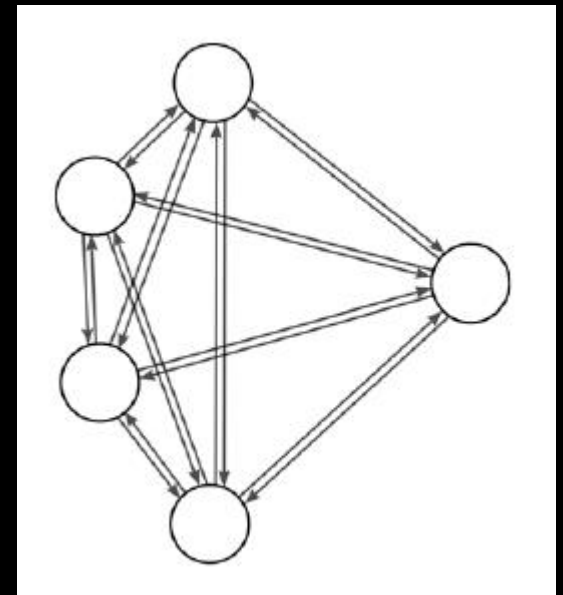
ANN

Neurónové siete - siete pospájaných uzlov

- informácia sa presúva medzi uzlami
- uzol vyšle informáciu na základe vstupu



Orientovaný acyklický graf
Dopredná sieť



Spätná sieť

ANN

Postup:

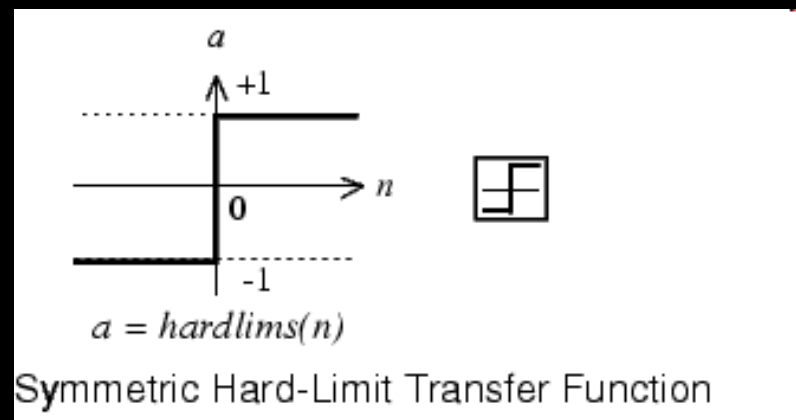
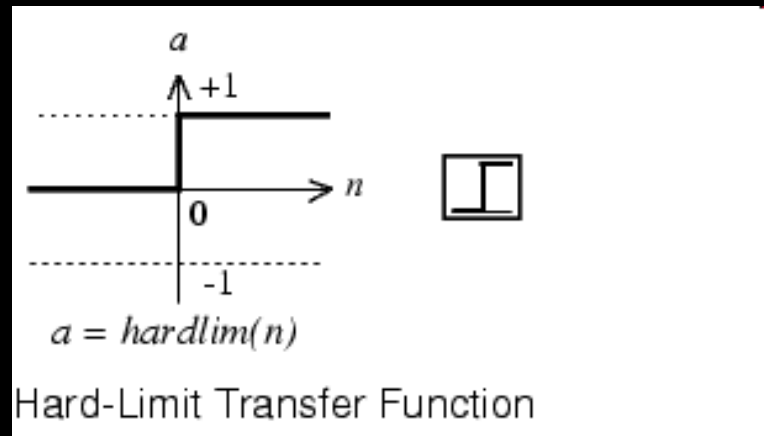
- výstupy z uzlov predchádzajúcej vrstvy sa ovážia a spočítajú
- na túto sumu sa aplikuje nelineárna aktivačná funkcia φ
- tento výstup sa pošle ďalej



Aktivačná funkcia

prahová

$$\phi(x) = \begin{cases} 1 & \text{ak } x \geq 0, \\ 0 & \text{inak.} \end{cases}$$



rýchla a plynulá zmena

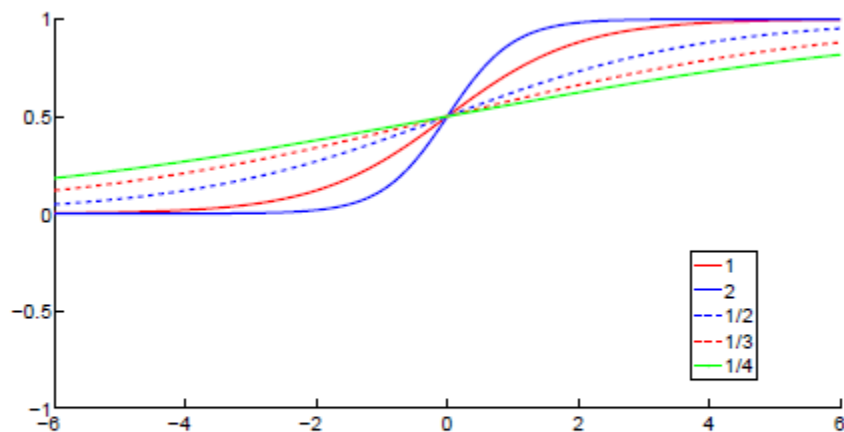
logistická funkcia

(pre $\alpha = 1$ nazývaná sigmoida)

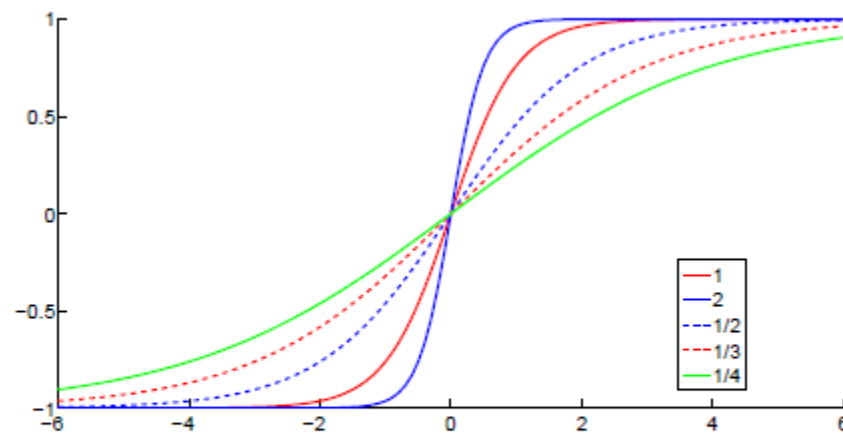
$$\phi(\alpha, x) = \frac{1}{1 + e^{-\alpha x}},$$

hyperbolický tangens

$$\phi(\beta, x) = \tanh(\beta x) = \frac{e^{\beta x} - e^{-\beta x}}{e^{\beta x} + e^{-\beta x}}.$$

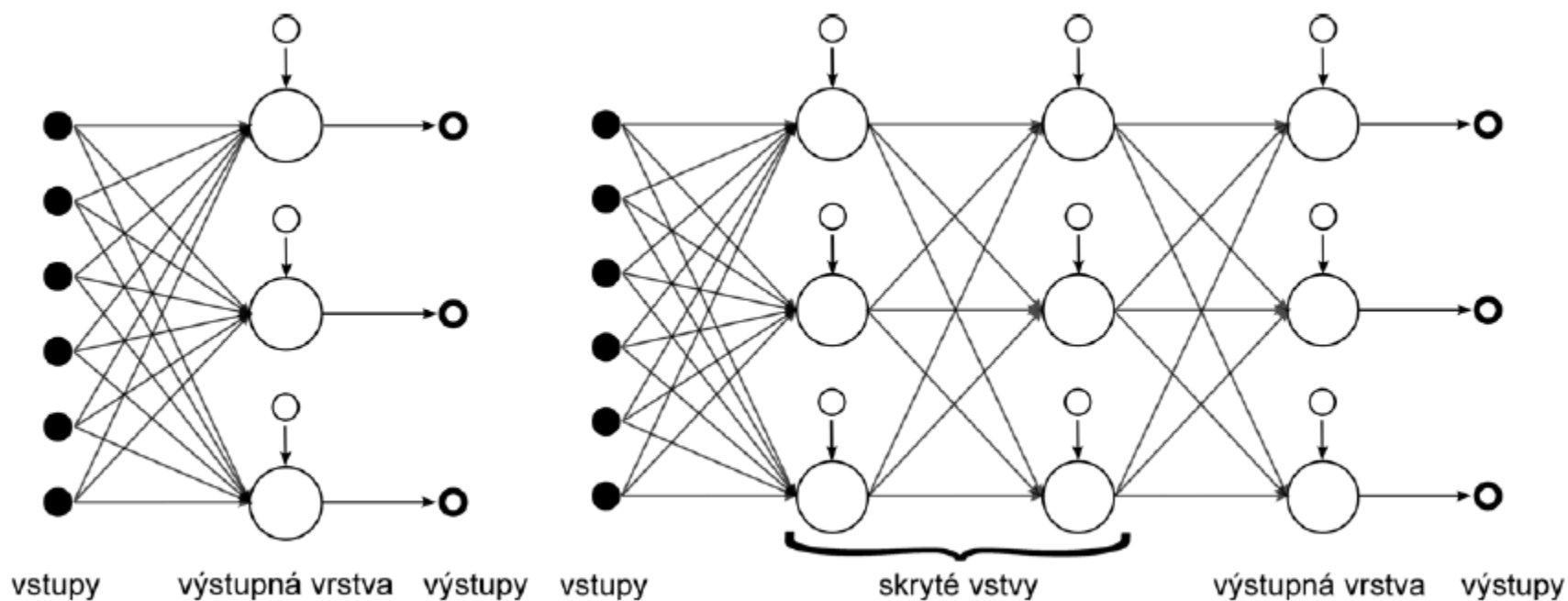


(a) Logistická funkcia pre rôzne hodnoty parametra α .



(b) Hyperbolický tangens pre rôzne hodnoty parametra β .

Viacvrstvý perceptrón



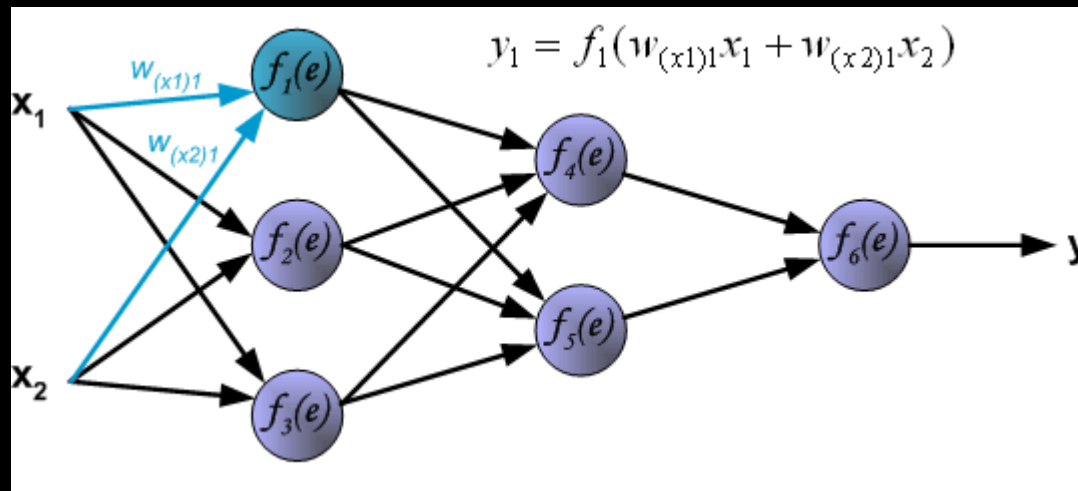
Učenie neurónovej siete

určenie váhovej matice (váhových vektorov v jednotlivých vrstvách)

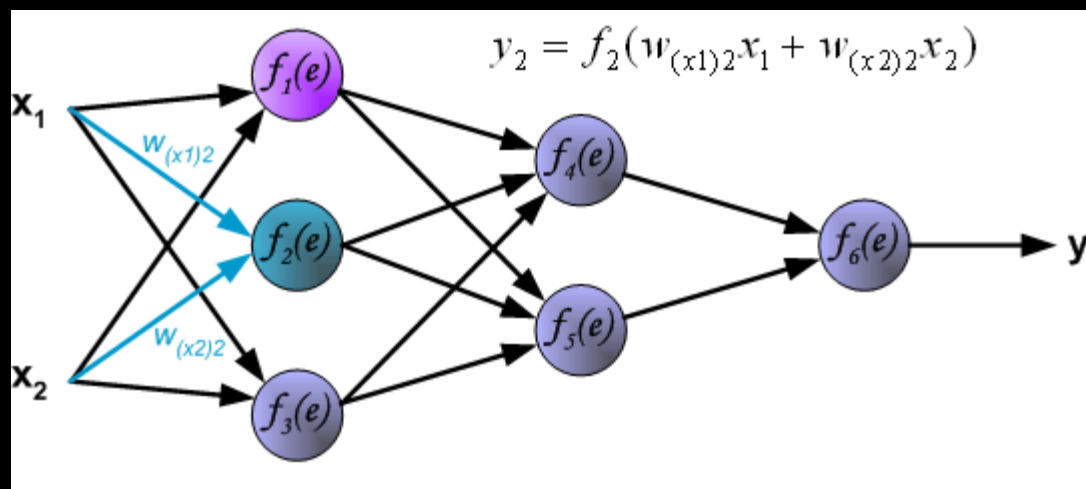
1. inicializácia náhodnej váhovej matice
2. vektory príznačkov z trénovacej množiny postupne prichádzajú na vstup a „prebublajú“ sieťou na koniec
3. výstupný vektor sa porovná s požadovanou klasifikáciou a určí sa hodnota kritéria E
4. chyba sa spätne šíri a upravujú sa hodnoty váhovej matice

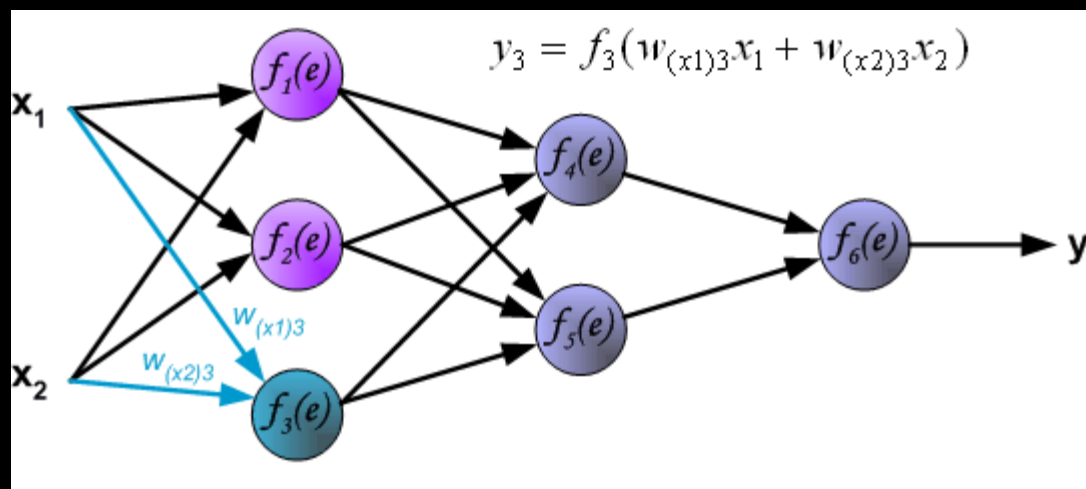
$$W \leftarrow W + \Delta W,$$

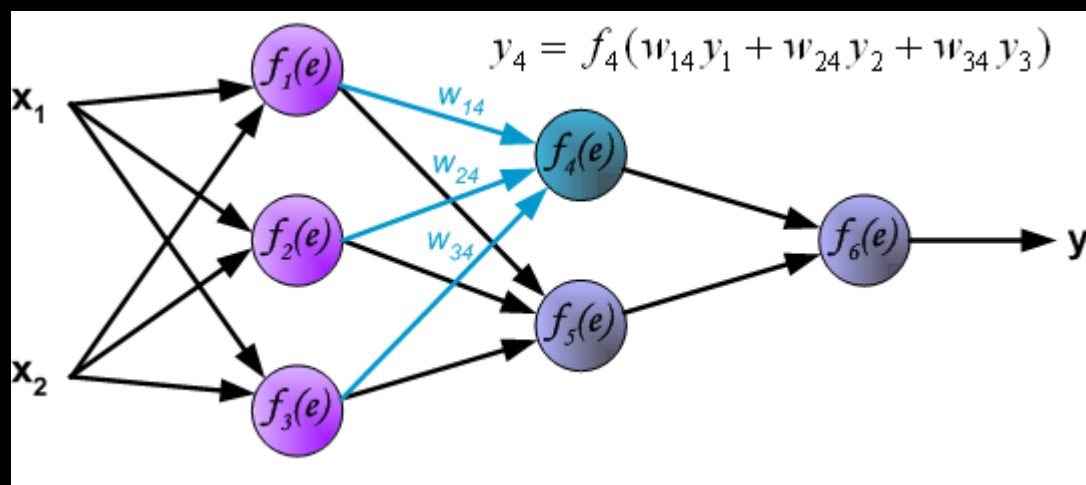
$$\Delta W = -\eta \nabla E|_W.$$

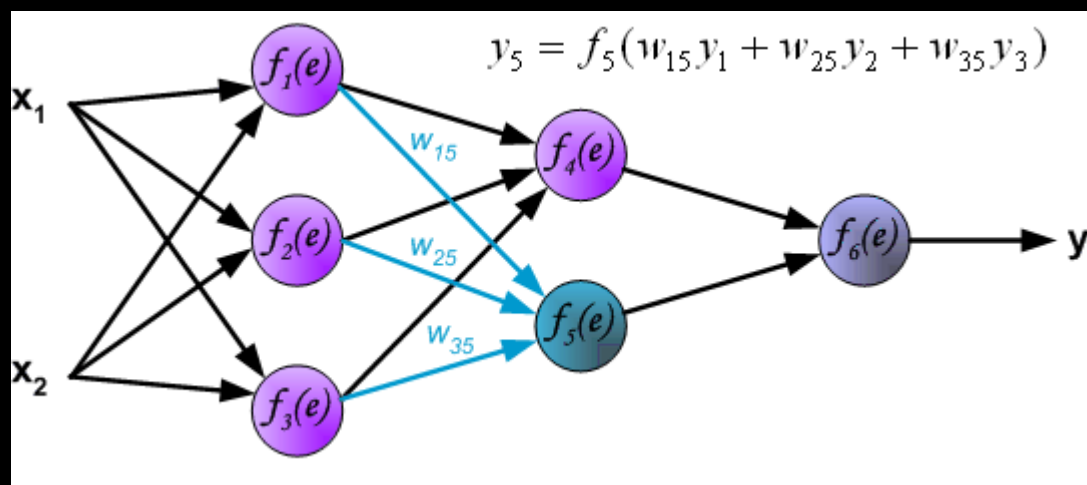


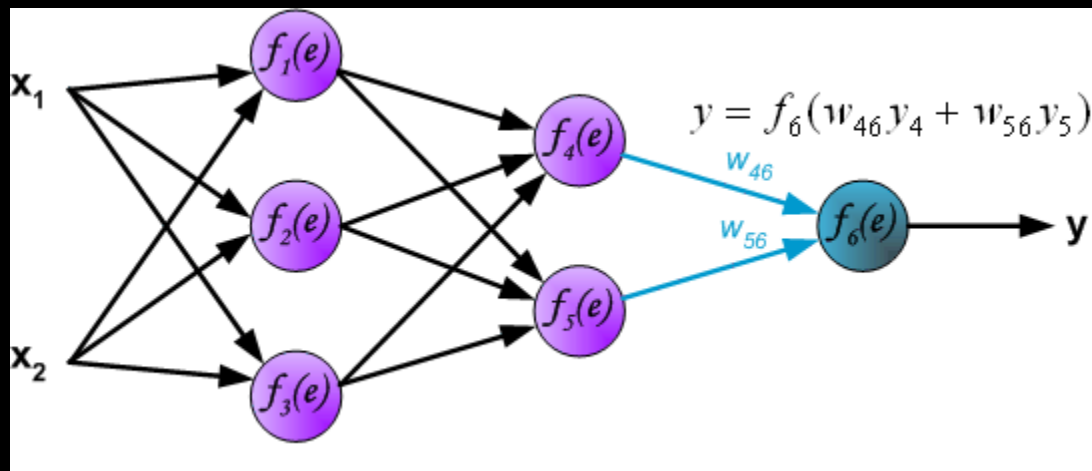
http://galaxy.agh.edu.pl/~vlsi/AI/backp_t_en/backprop.html

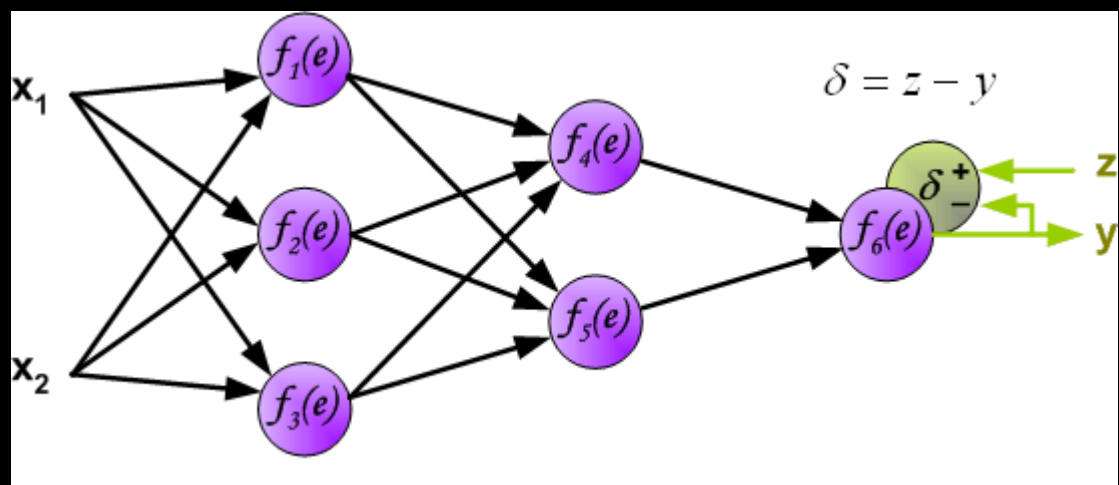


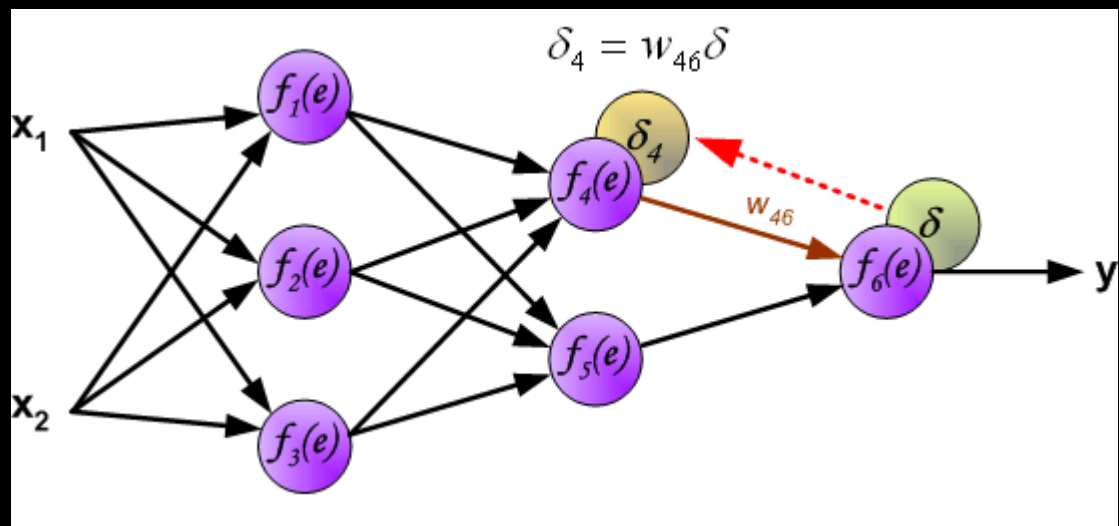


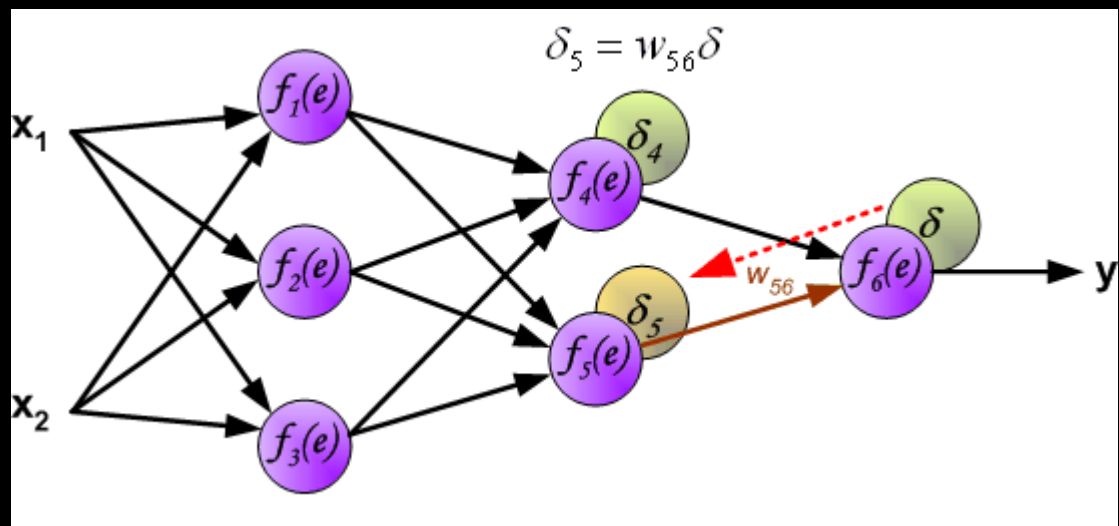


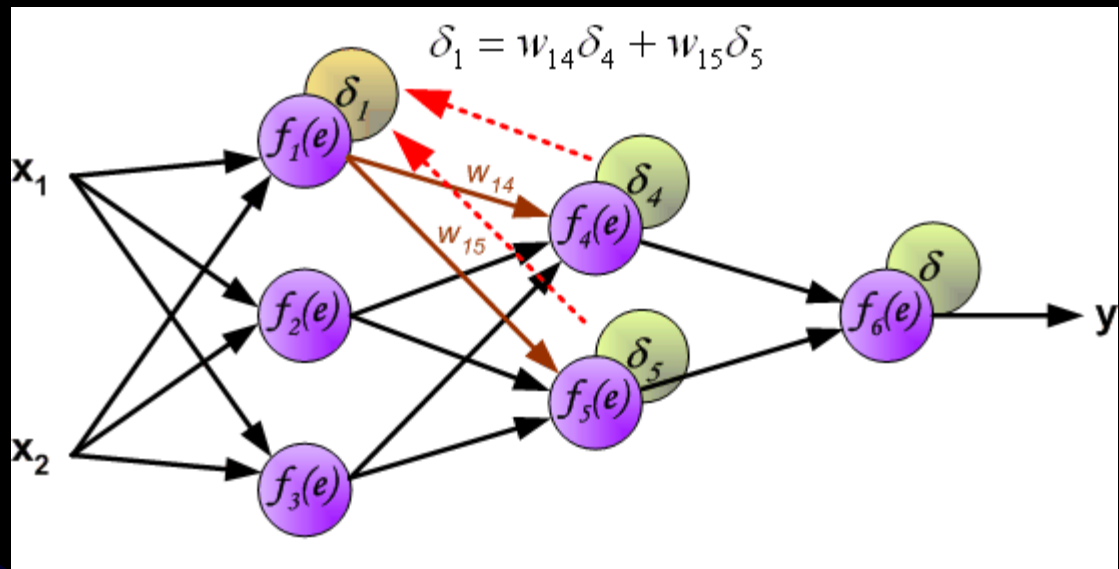


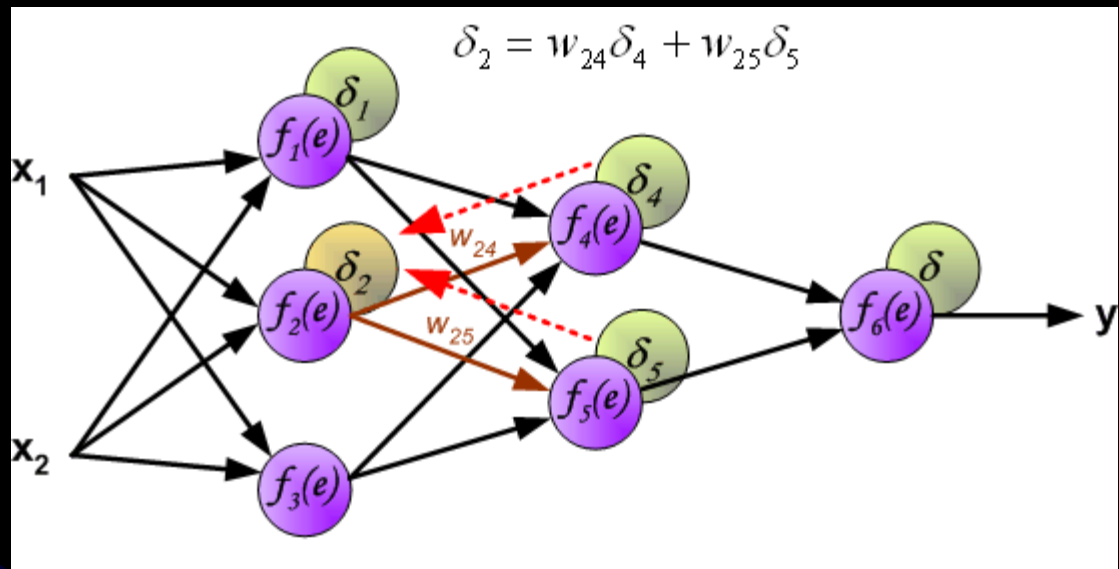


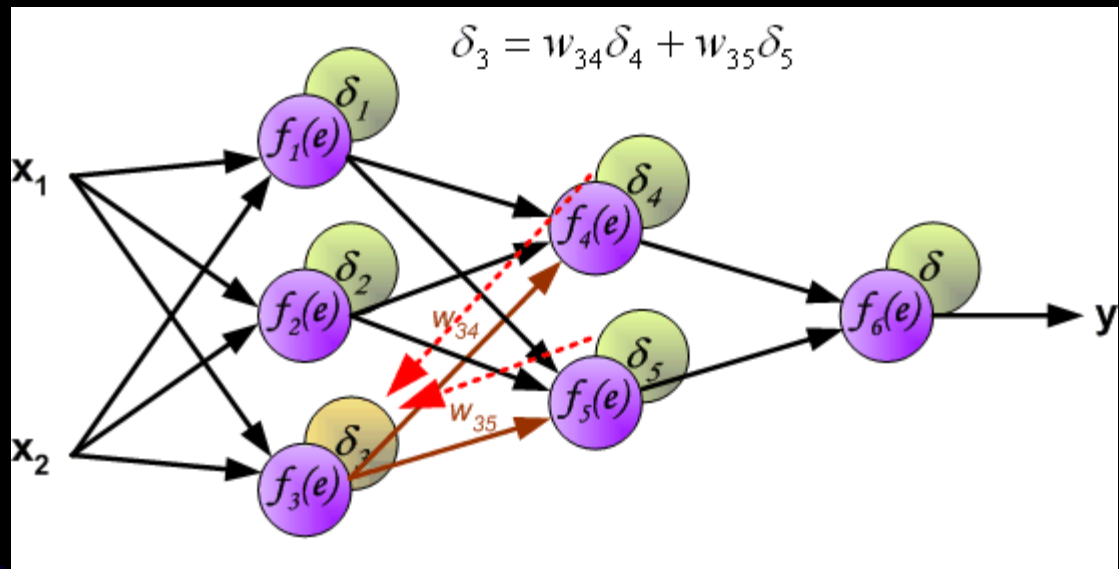


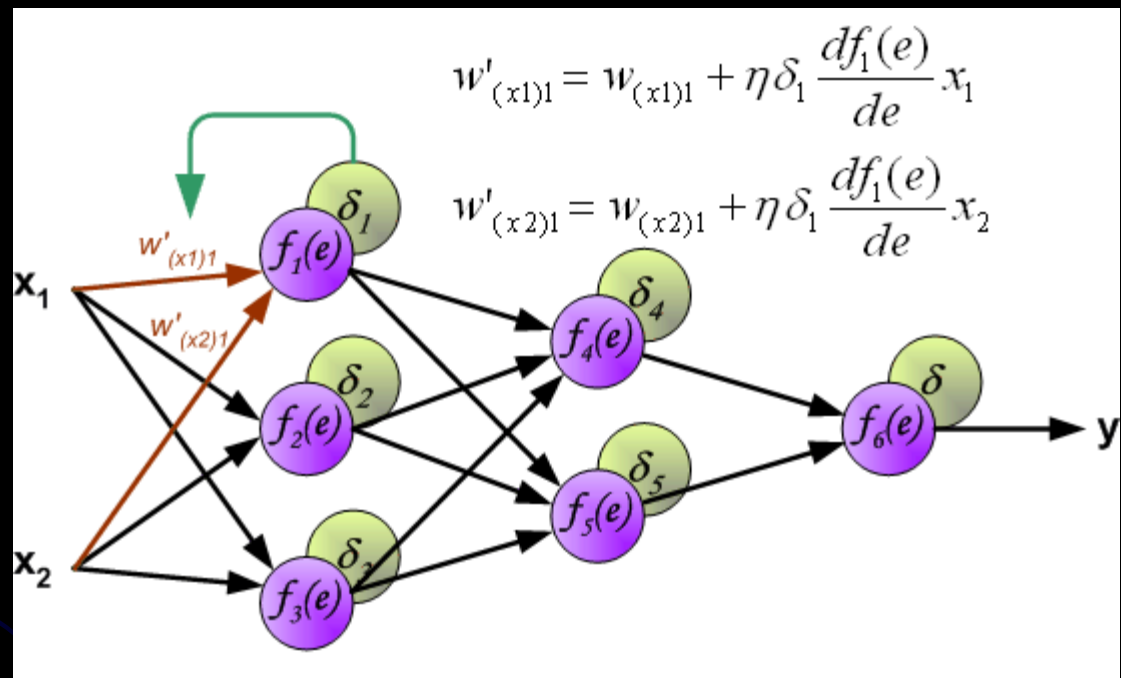






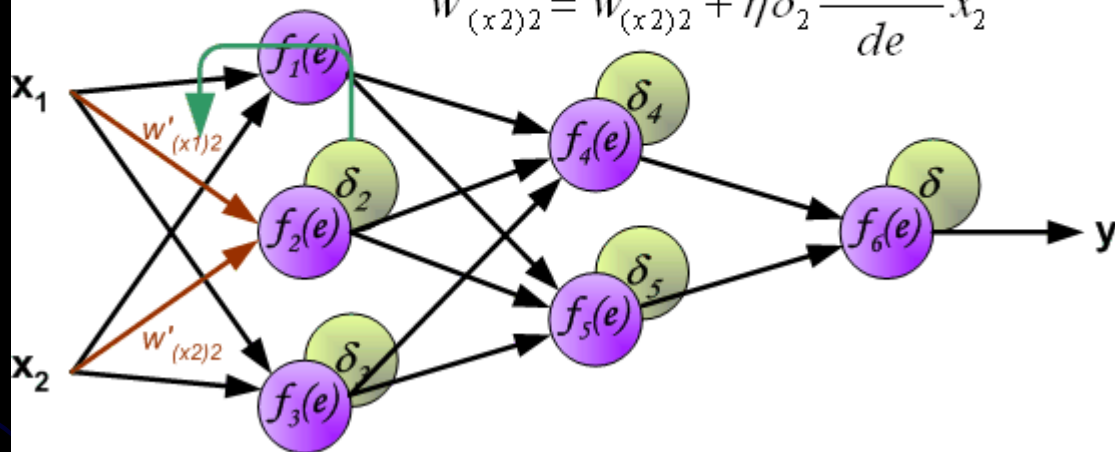






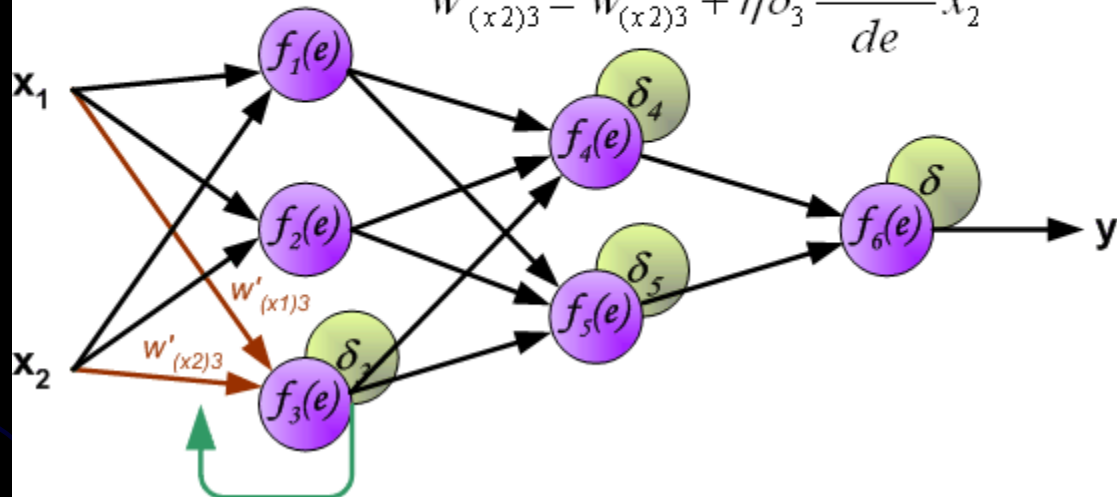
$$w'_{(x1)2} = w_{(x1)2} + \eta \delta_2 \frac{df_2(e)}{de} x_1$$

$$w'_{(x2)2} = w_{(x2)2} + \eta \delta_2 \frac{df_2(e)}{de} x_2$$



$$w'_{(x1)3} = w_{(x1)3} + \eta \delta_3 \frac{df_3(e)}{de} x_1$$

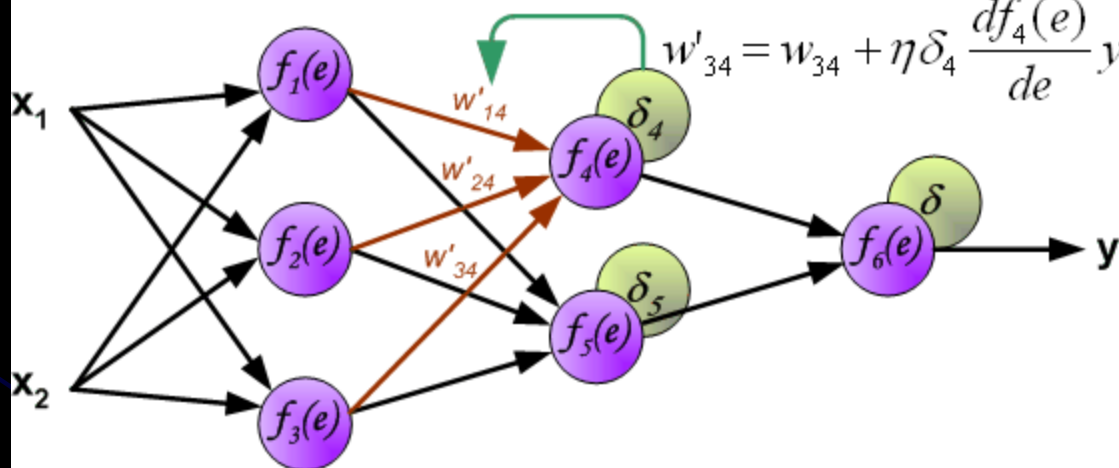
$$w'_{(x2)3} = w_{(x2)3} + \eta \delta_3 \frac{df_3(e)}{de} x_2$$



$$w'_{14} = w_{14} + \eta \delta_4 \frac{df_4(e)}{de} y_1$$

$$w'_{24} = w_{24} + \eta \delta_4 \frac{df_4(e)}{de} y_2$$

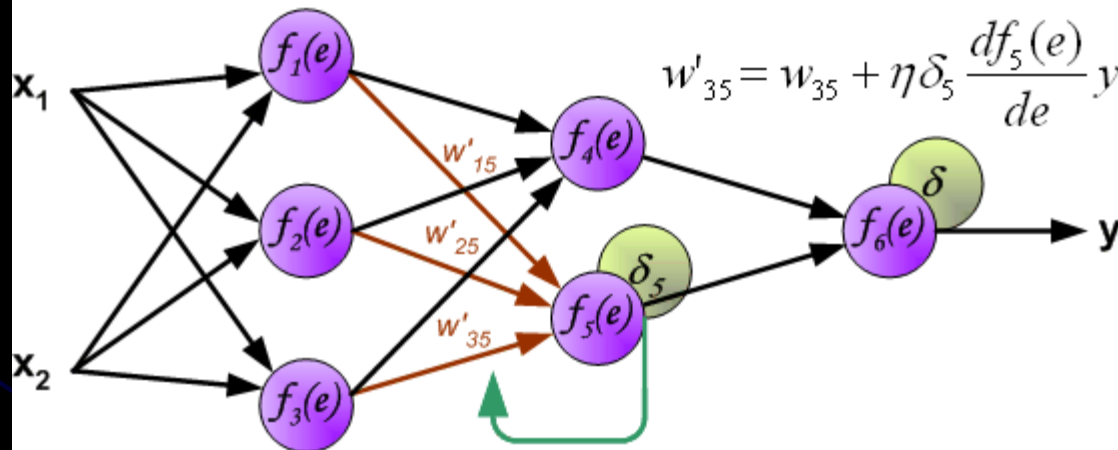
$$w'_{34} = w_{34} + \eta \delta_4 \frac{df_4(e)}{de} y_3$$

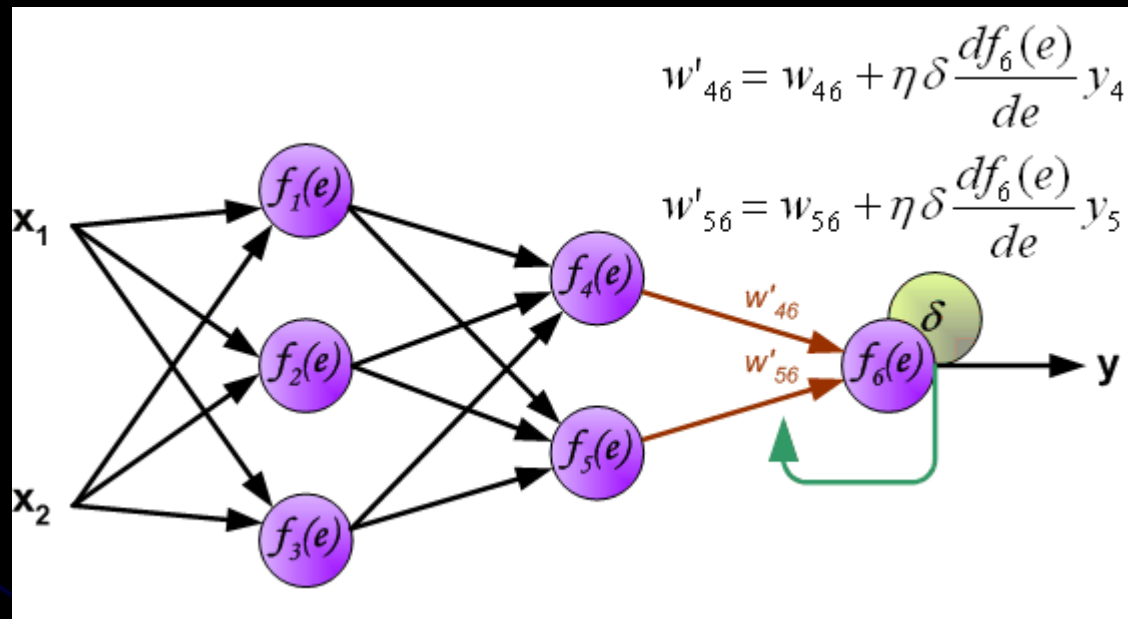


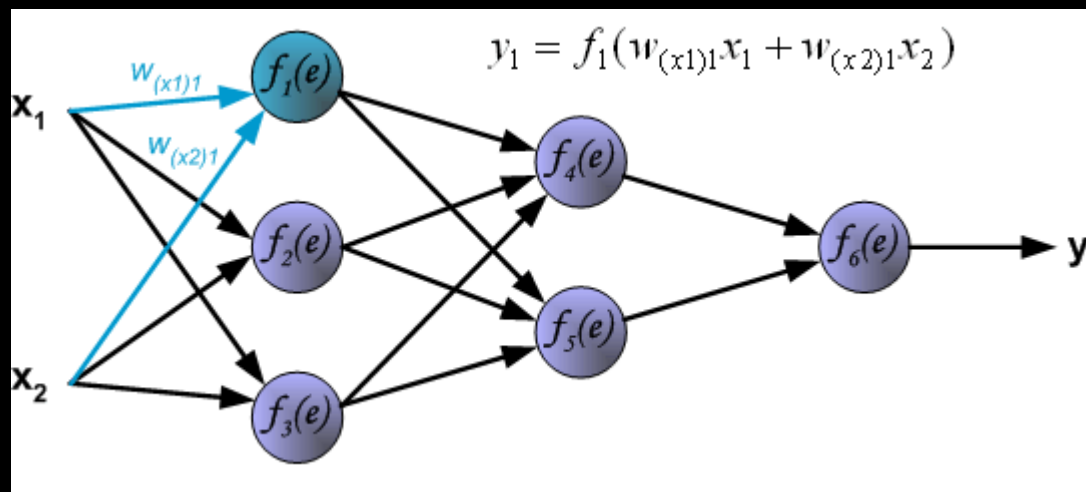
$$w'_{15} = w_{15} + \eta \delta_5 \frac{df_5(e)}{de} y_1$$

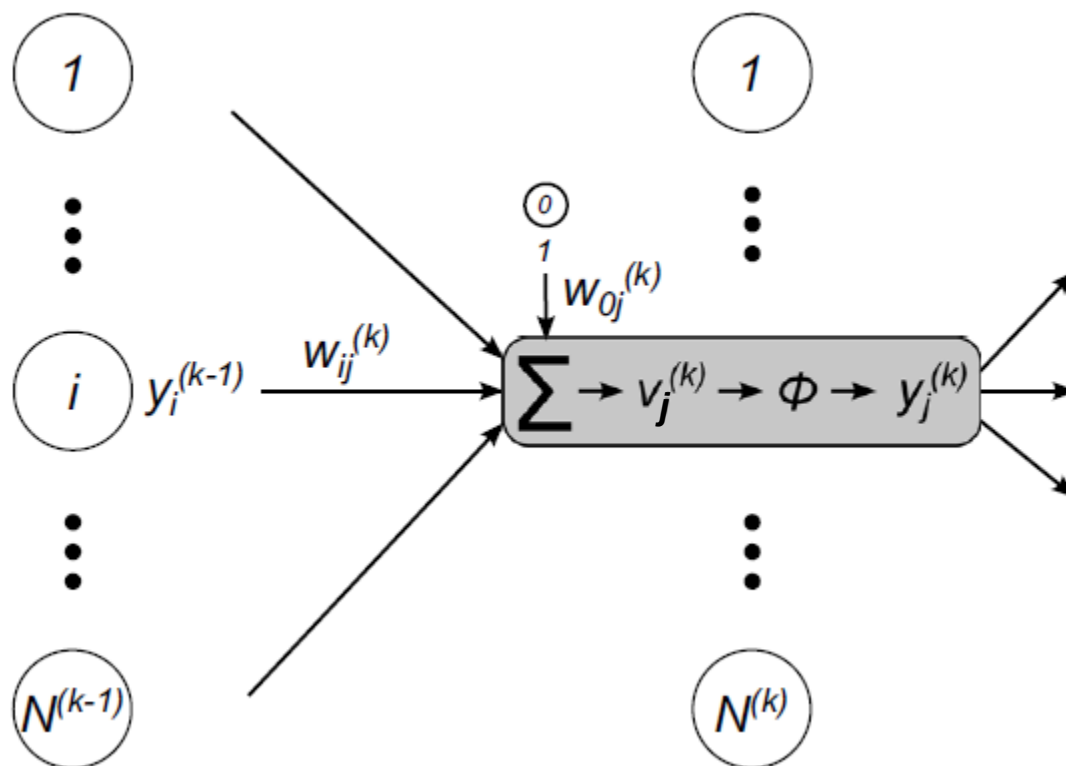
$$w'_{25} = w_{25} + \eta \delta_5 \frac{df_5(e)}{de} y_2$$

$$w'_{35} = w_{35} + \eta \delta_5 \frac{df_5(e)}{de} y_3$$









$$v_j^{(k)} = \sum_{i=1}^{N^{(k-1)}} w_{ij}^{(k)} y_i^{(k-1)} + w_{0j}^{(k)},$$

$$y_j^{(k)} = \phi^{(k)} \left(v_j^{(k)} \right)$$

$$\Delta w_{ij}^{(k)} = -\eta \frac{\partial E}{\partial w_{ij}^{(k)}}.$$

$$\mathbf{W} \leftarrow \mathbf{W} + \Delta \mathbf{W},$$

$$\Delta \mathbf{W} = -\eta \nabla E|_{\mathbf{W}}.$$

$$\Delta w_{ij}^{(k)} = -\eta \frac{\partial E}{\partial v_j^{(k)}} \frac{\partial v_j^{(k)}}{\partial w_{ij}^{(k)}}$$

$$\frac{\partial v_j^{(k)}}{\partial w_{ij}^{(k)}} = y_i^{(k-1)}$$

$$\delta_j^{(k)} = -\frac{\partial E}{\partial v_j^{(k)}} = -\frac{\partial E}{\partial y_j^{(k)}} \frac{\partial y_j^{(k)}}{\partial v_j^{(k)}}$$

$$\Delta w_{ij}^{(k)} = \eta \delta_j^{(k)} y_i^{(k-1)}$$

Šírenie chyby vo výstupnej vrstve

stredná kvadratická chyba (MSE)

$$E(\mathbf{W}) = \frac{1}{2} \sum_{j=1}^M (y_j^{(m)} - o_j)^2$$

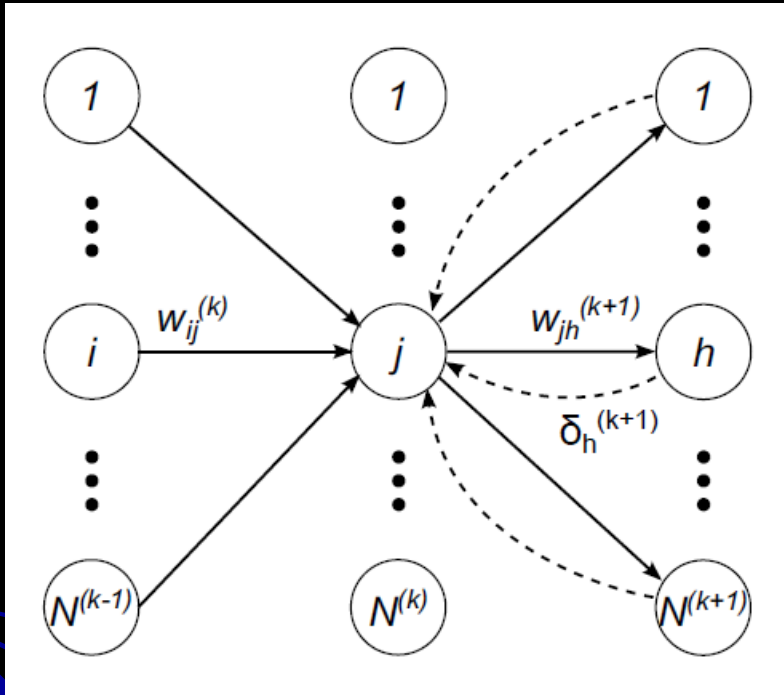
$$\delta_j^{(m)} = -\frac{\partial E}{\partial v_j^{(m)}} = -\frac{\partial E}{\partial y_j^{(m)}} \frac{\partial y_j^{(m)}}{\partial v_j^{(m)}}$$

$$\frac{\partial E}{\partial y_j^{(m)}} = y_j^{(m)} - o_j$$

$$\frac{\partial y_j^{(m)}}{\partial v_j^{(m)}} = \phi'(v_j^{(m)})$$

$$\delta_j^{(m)} = (o_j - y_j^{(m)})\phi'(v_j^{(m)})$$

Šírenie chyby v skrytých vrstvách



$$\frac{\partial E}{\partial y_j^{(k)}} = \sum_{h=1}^{N^{(k+1)}} \frac{\partial E}{\partial v_h^{(k+1)}} \frac{\partial v_h^{(k+1)}}{\partial y_j^{(k)}} =$$

$$= - \sum_{h=1}^{N^{(k+1)}} \delta_h^{(k+1)} w_{jh}^{(k+1)}$$

$$\frac{\partial y_j^{(k)}}{\partial v_j^{(k)}} = \phi'(v_j^{(k)})$$

$$\delta_j^{(k)} = \phi'(v_j^{(k)}) \sum_{h=1}^{N^{(k+1)}} \delta_h^{(k+1)} w_{jh}^{(k+1)}$$

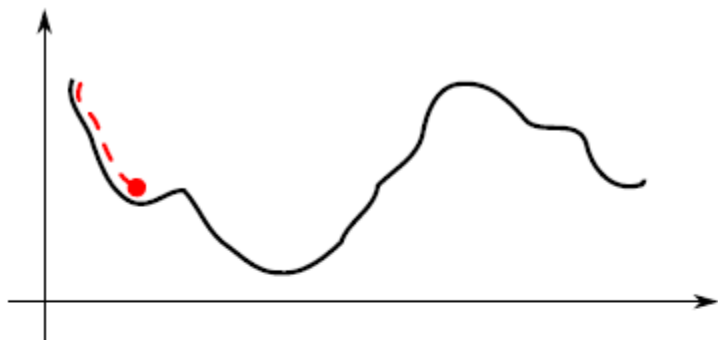
$$\delta_j^{(m)} = (o_j - y_j^{(m)})\phi'(v_j^{(m)})$$

$$\delta_j^{(k)} = \phi'(v_j^{(k)}) \sum_{h=1}^{N^{(k+1)}} \delta_h^{(k+1)} w_{jh}^{(k+1)}$$

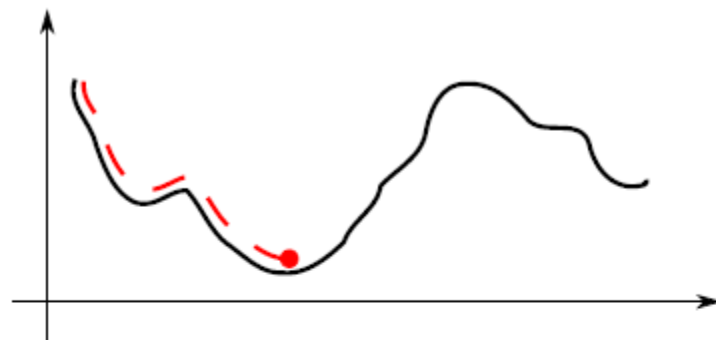
$$\mathbf{W} \leftarrow \mathbf{W} + \Delta \mathbf{W},$$

$$\Delta w_{ij}^{(k)} = \eta \delta_j^{(k)} y_i^{(k-1)}$$

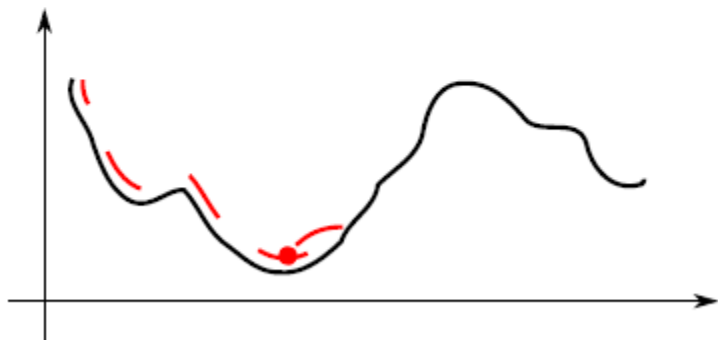
Rýchlosť učenia sa neurónovej siete



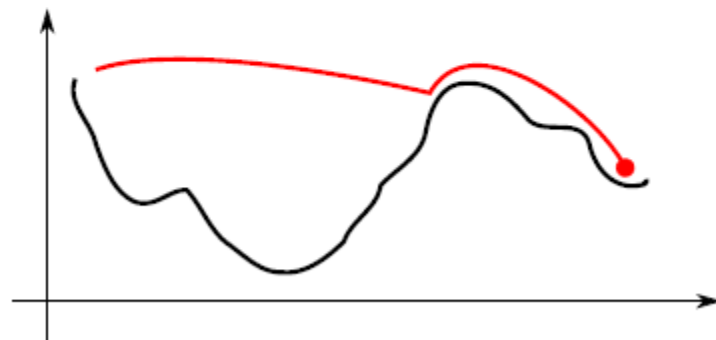
(a) Nízka rýchlosť učenia sa.



(b) Optimálna rýchlosť učenia sa.



(c) Vysoká rýchlosť učenia sa.



(d) Veľmi vysoká rýchlosť učenia sa.

Nízka rýchlosť učenia – pomalý tréning, môže skončiť v lokálnom minime

Optimálna rýchlosť učenia – pomalý tréning, ale bez oscilácií, dosiahne globálne optimum

- Vysoká rýchlosť učenia – rýchlejší tréning, ale môže oscilovať, prípadne neskončiť v globálnom optime

Heuristika – premenlivá rýchlosť učenia:

Ak nová chyba prevýši starú o stanovený prah, ideme zlým smerom. Preto sa nové váhy zahodia a rýchlosť učenia sa sa zníži.

Malé zvýšenie chyby povolíme kvôli možnosti dostať sa z lokálneho minima. V prípade, že nová chyba je vyššia ako stará o menej než stanovený prah, váhy sa upraví.

Ak by nová chyba bola nižšia než stará, tak smerujeme k minimu, môžeme upraviť váhy a zvýšiť rýchlosť učenia sa.

Ukončenie učenia sa neurónovej siete

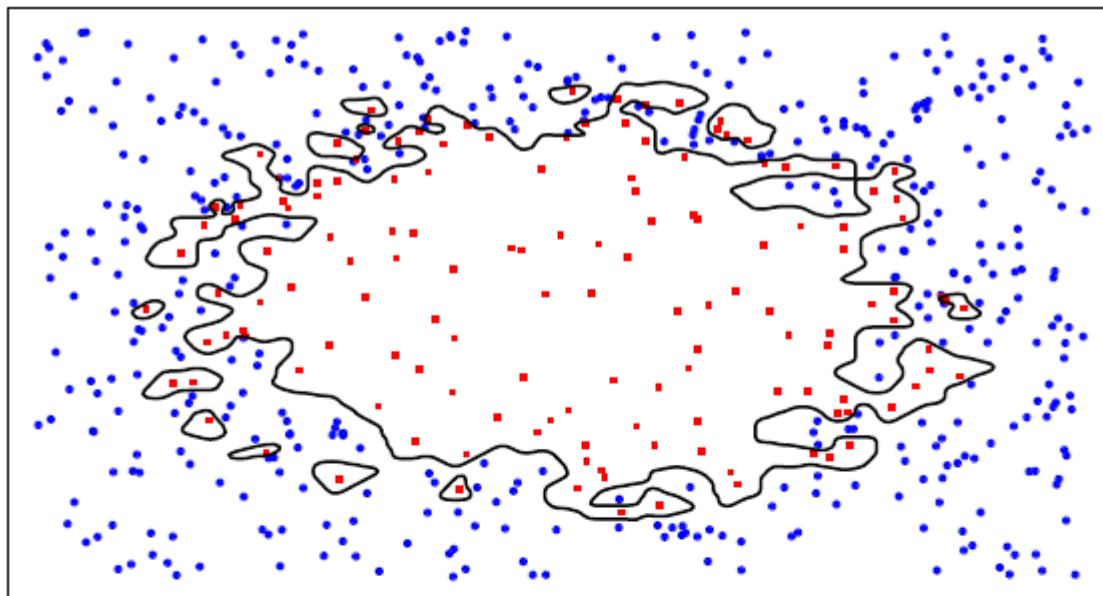
ideálne – keď dosiahneme globálne minimum chybovej funkcie

Problém

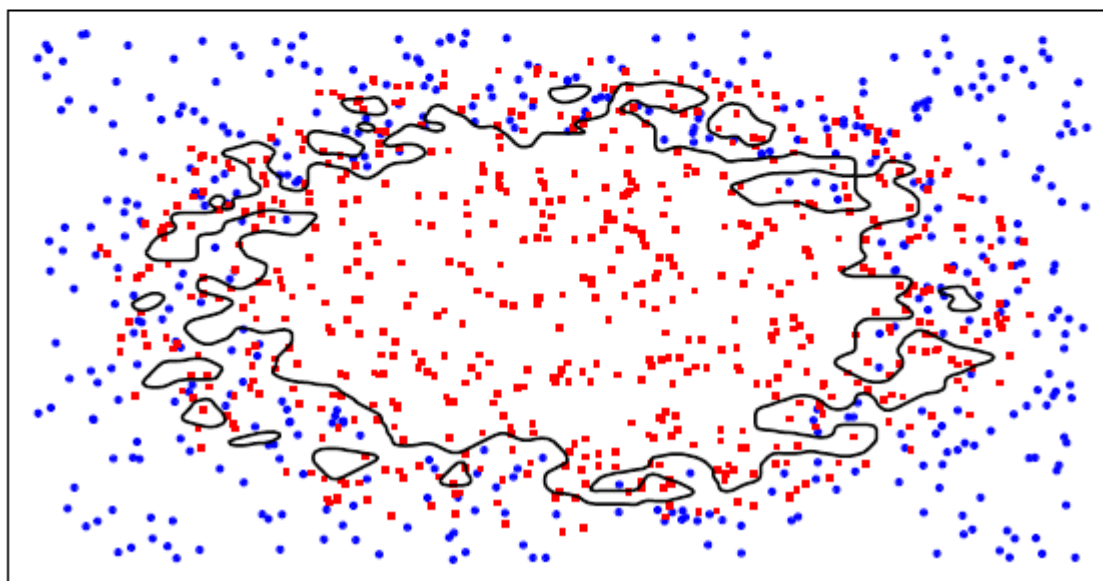
Preučenie siete – malá chyba na tréningových dátach, ale veľká na testovacích

Riešenie

- určíme tréningovú a testovaciu množinu
- ukončíme tréning, keď sa chyba v testovacej množine začne zvyšovať



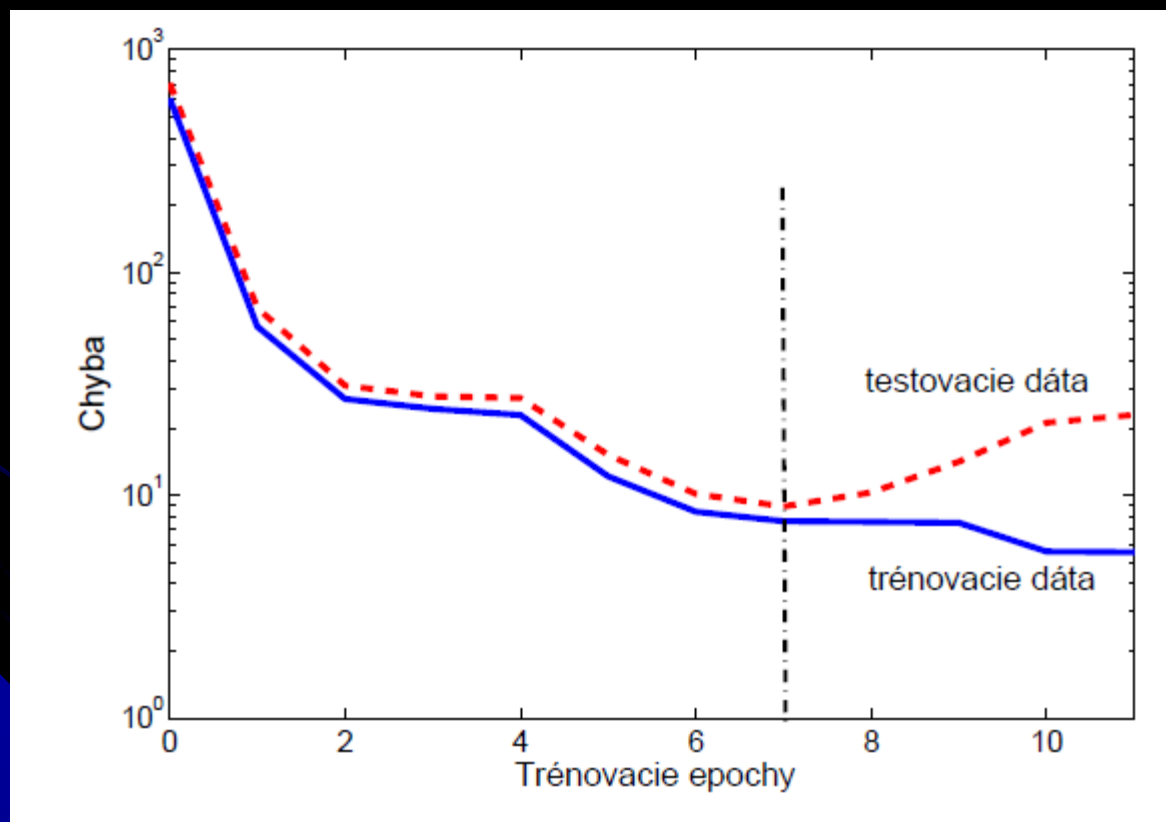
(a) Trénovacie dáta.



(b) Testovacie dáta.

Riešenie

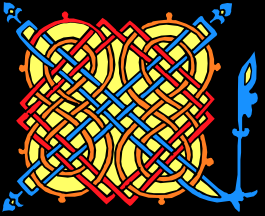
určíme trénováciu a testovaciu množinu
ukončíme tréning, keď sa chyba v testovacej množine
začne zvyšovať



Aplikácie

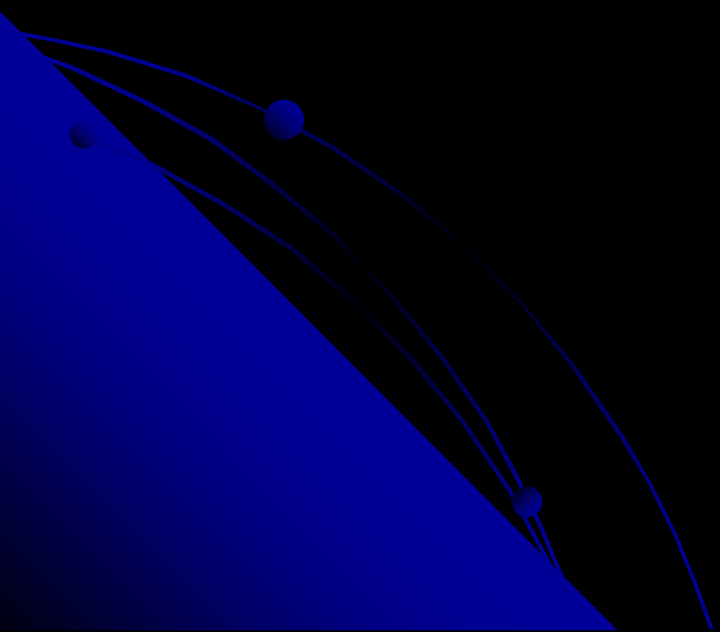
<http://tralvex.com/pub/nap/>





Self-Organizing Maps

Samoorganizujúce sa mapy

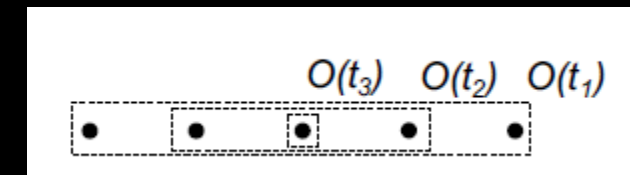
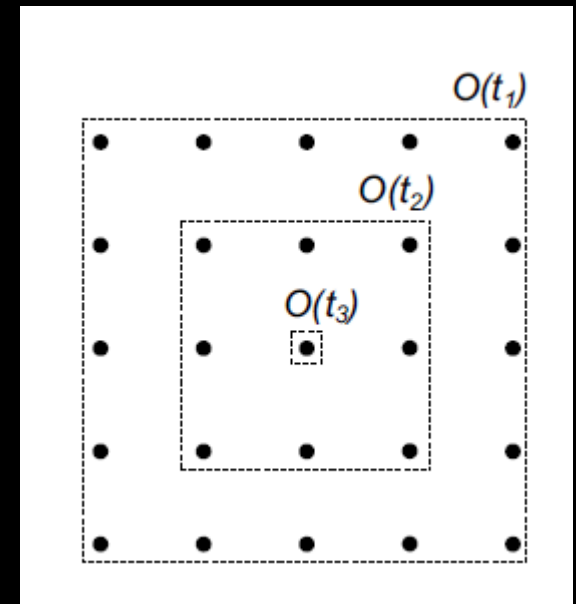


SOM

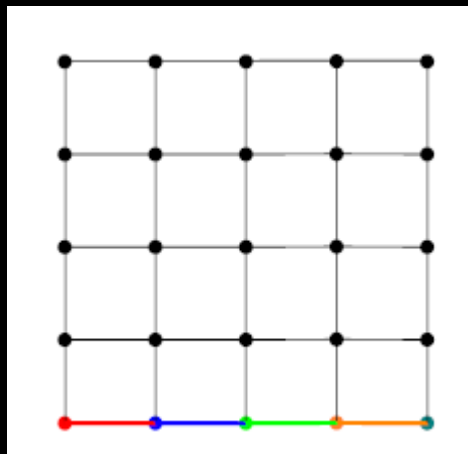
Samoorganizujúce sa mapy (Kohonenove siete)

Neuróny sú umiestnené v mriežke
Vstupné dáta
sú spojené s každým neurónom

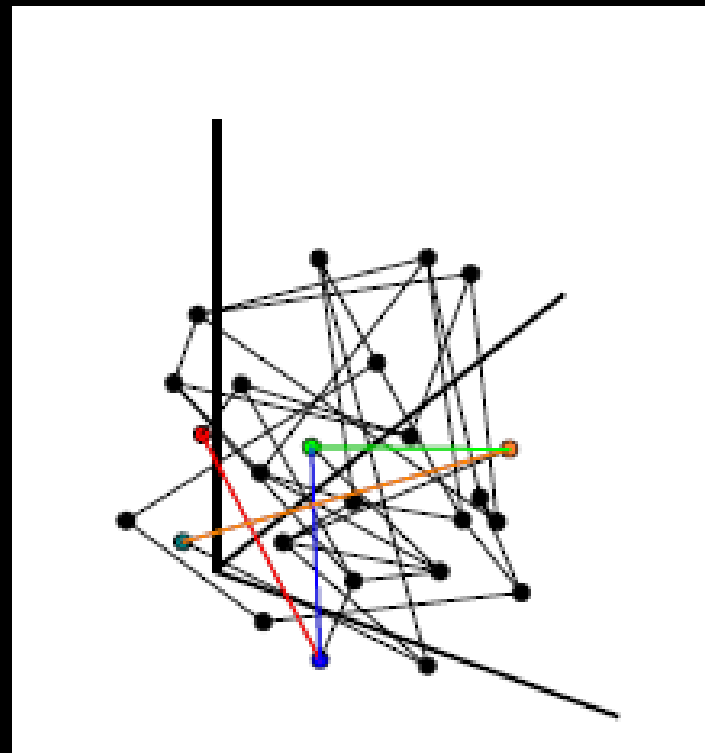
Neriadená metóda



Dva priestory SOM



Mriežka SOM,
topologické susedstvo



Priestor váh

Učenie SOM

Inicializácia

Učenie siete:

1. Vstup vektora príznakov
2. Určenie víťazného neurónu
3. Víťaz berie všetko

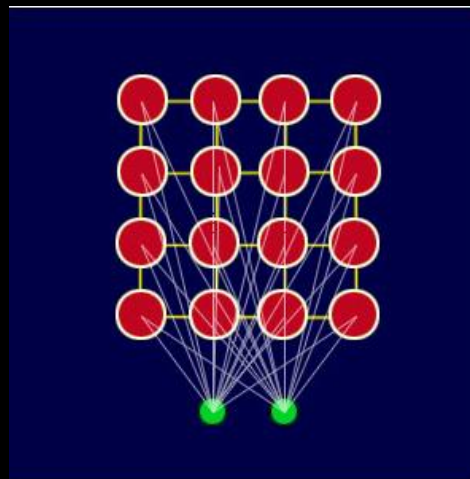
1. Krok

Inicializácia a normalizácia váhových vektorov:

$$\bar{w}_i = \frac{w_i}{\|w_i\|}, \quad 1 \leq i \leq c$$

2. Krok

Vstup vektora príznačov



3. krok

Určenie víťazného neurónu

$$i^* = \underset{i}{\operatorname{argmin}} \| \mathbf{x} - \mathbf{w}_i \|$$

hľadá sa neurón, ktorého váhový vektor je najbližšie k aktuálnemu vstupu v zmysle euklidovskej vzdialenosti
 c počet klasifikačných tried – počet neurónov v sieti

4. krok

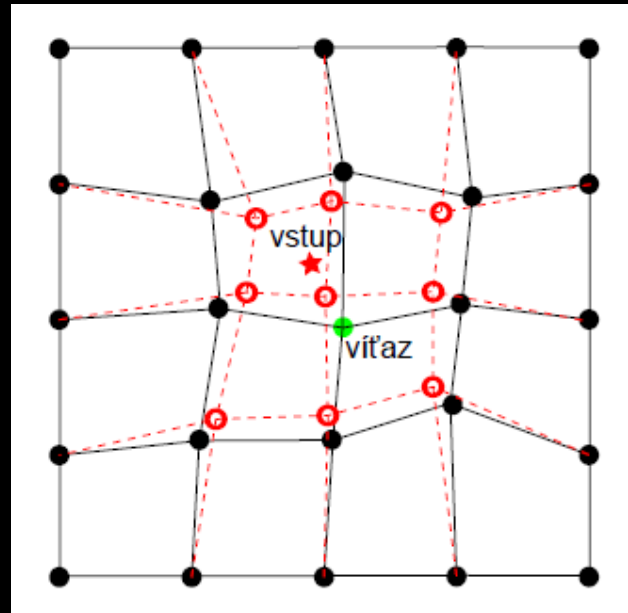
Víťaz berie všetko – víťaznému neurónu zmeníme váhu

Učenie: váhové vektory víťazného neurónu a jeho topologických susedov sa posunú smerom k aktuálnemu vstupu podľa vzťahu

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \alpha(t) O(i, i^*, t) \|\mathbf{x} - \mathbf{w}_i(t)\|$$

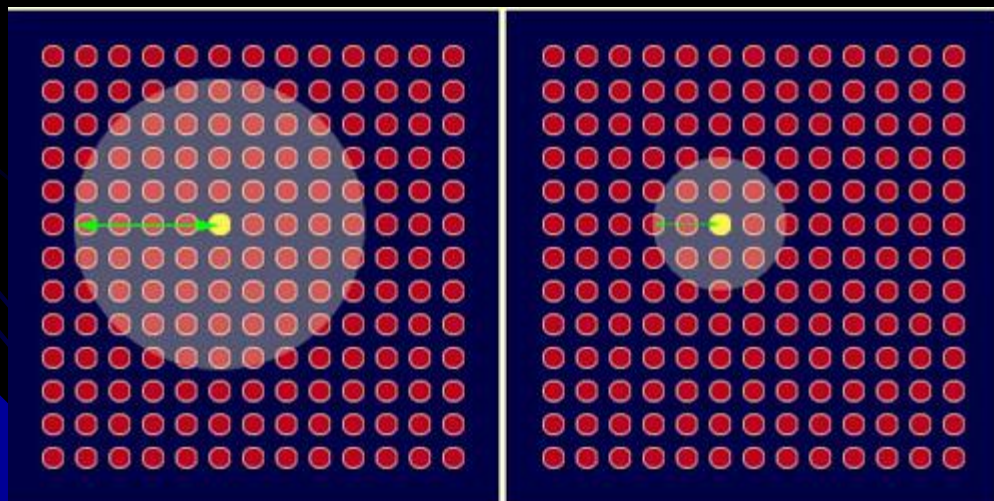
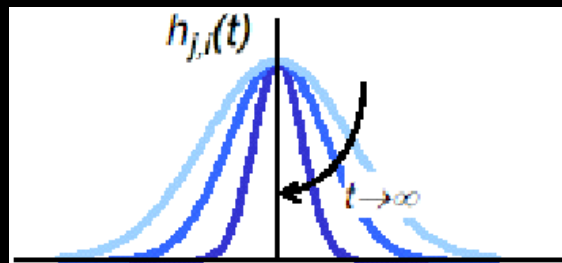
$\alpha(t) \in (0,1)$ Rýchlosť učenia (s časom klesá k nule), čím sa zabezpečí ukončenie procesu učenia.

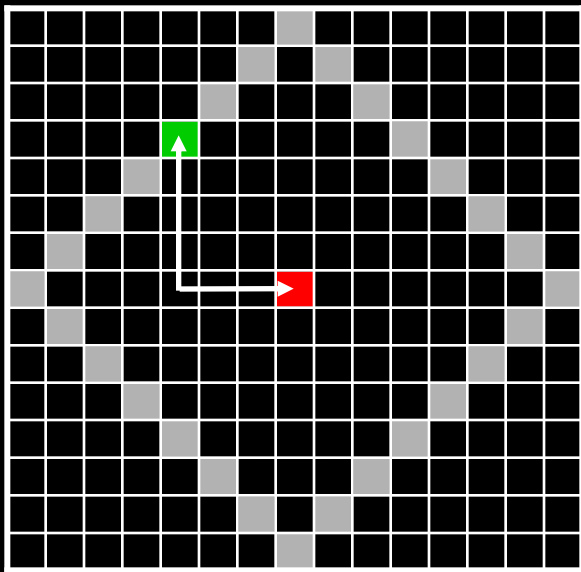
$O(i, i^*, t)$ Funkcia okolia - definuje rozsah kooperácie medzi neurónmi



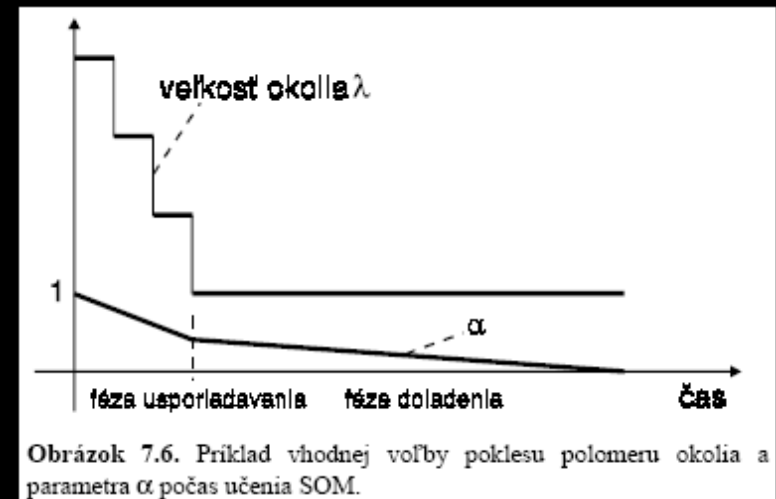
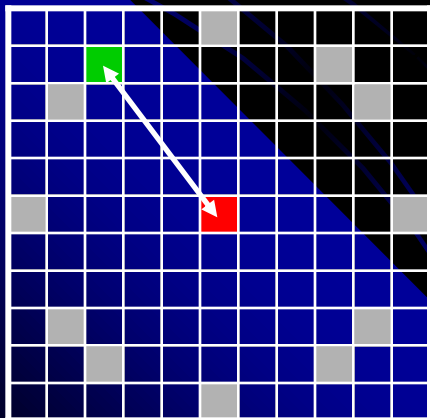
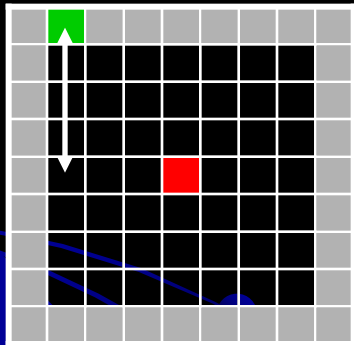
gaussovské okolie

$$O(i, i^*, t) = \exp \left(- \frac{(\mathbf{w}_{i^*}(t) - \mathbf{w}_i(t))^2}{\sigma} \right)$$



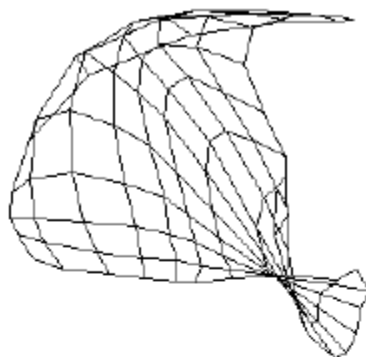


Rôzne metriky

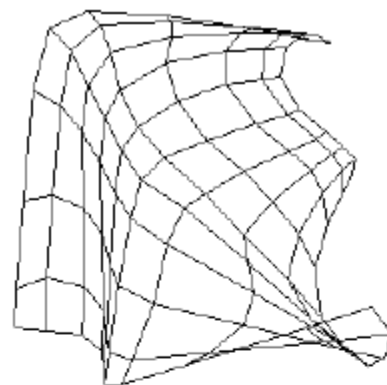




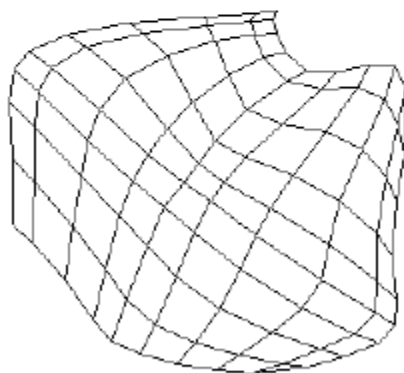
(a) Náhodné čísla z oblasti $\langle -0,5; 0,5 \rangle \times \langle -0,5; 0,5 \rangle$.



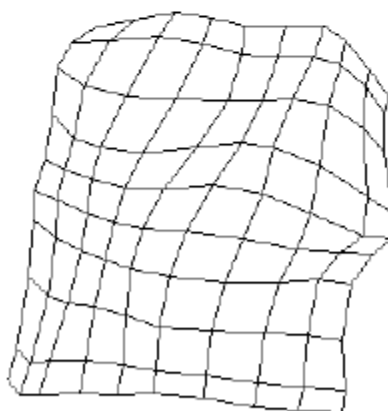
(b) Váhy po 100 iteráciách.



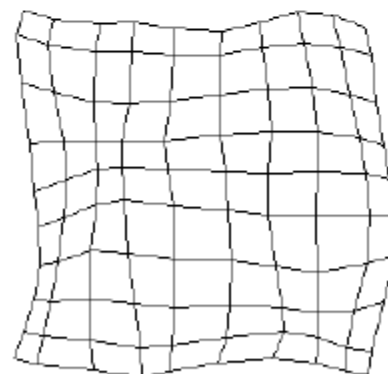
(c) Váhy po 200 iteráciách.



(d) Váhy po 600 iteráciách.

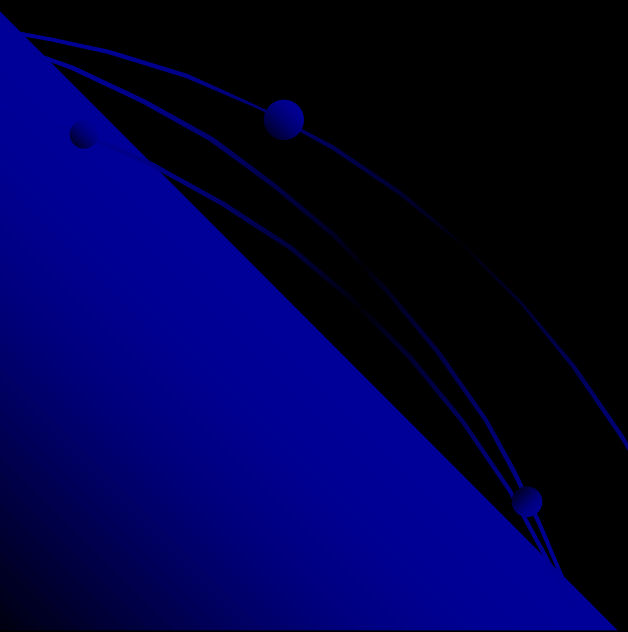
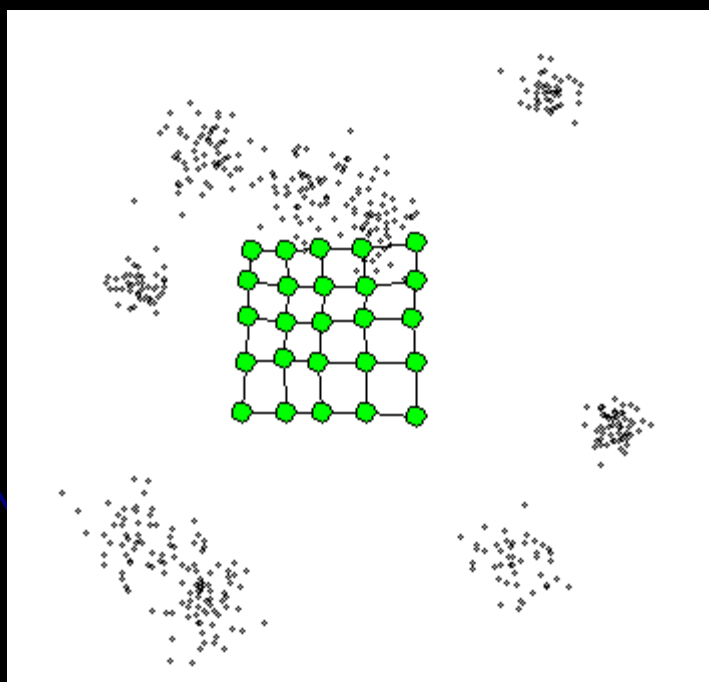
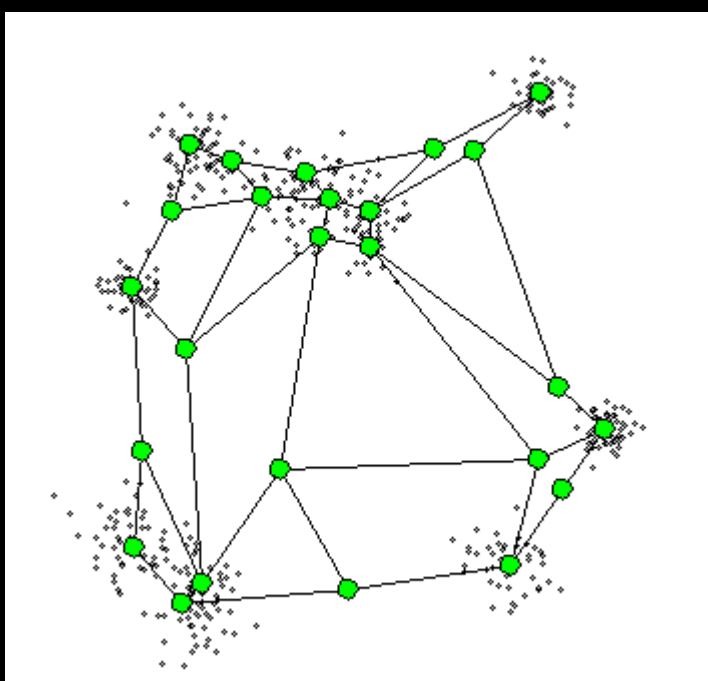
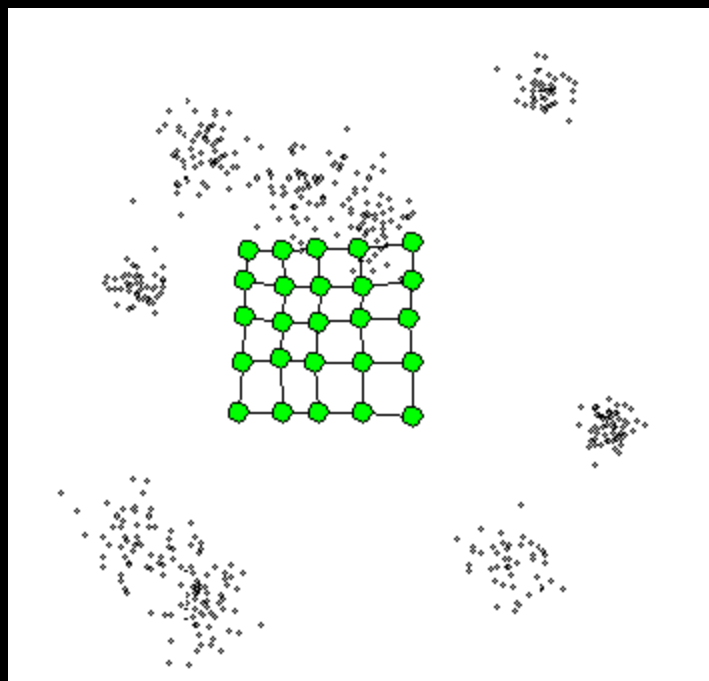


(e) Váhy po 3000 iteráciách.

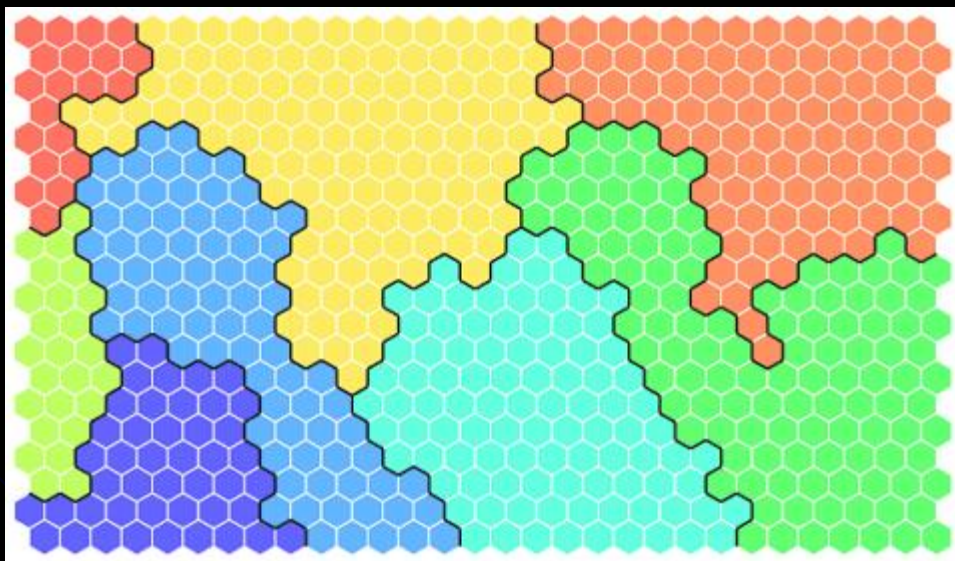


(f) Váhy po 7000 iteráciách.

Vstupnými vektormi sú rovnomerne rozdelené náhodné čísla z oblasti $\langle -1, 1 \rangle \times \langle -1, 1 \rangle$.

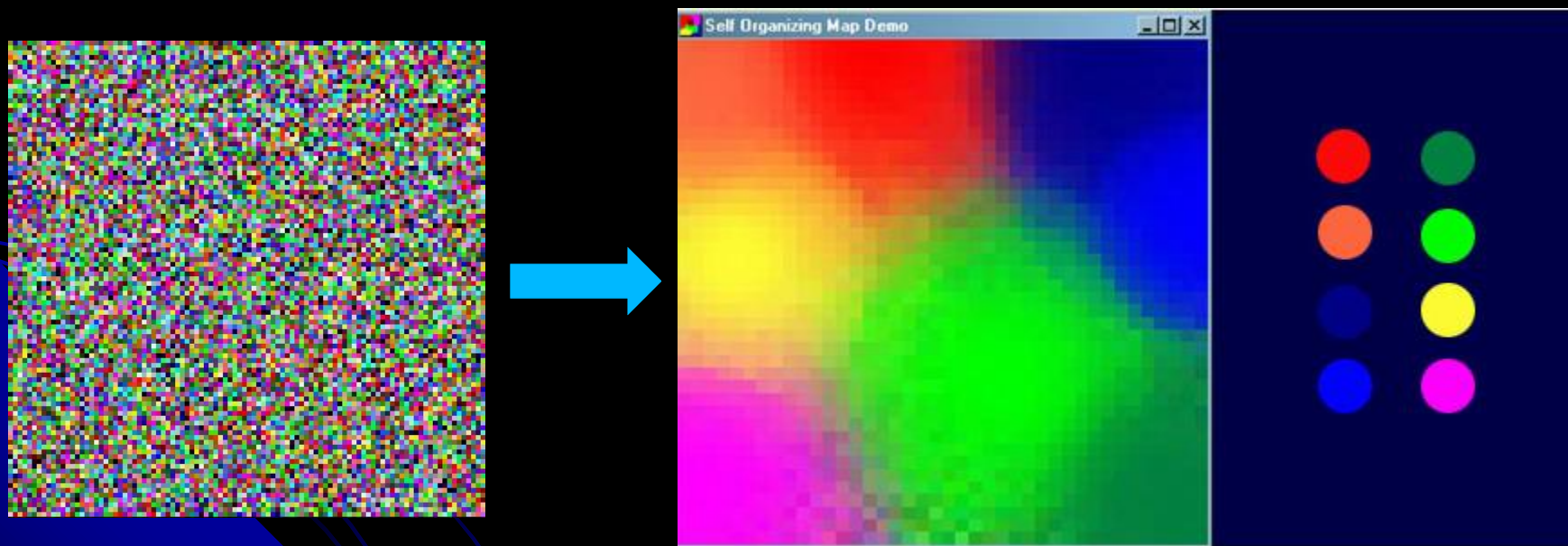


Klasifikácia



SOM Demo (ai-junkie.com):

3D farebný priestor reprezentovaný 2D SOM



SOM – mapa chudoby

Príklad z Helsinki University of Technology
dáta z World Bank statistics of countries in 1992

39 rozmerný vektor príznakov popisujúci kvalitu života v krajinách – zdravie, výživa, vzdelanie, ...

krajiny s podobnými hodnotami budú reprezentované blízkymi neurónmi

