

# Učenie neurónových sietí

# TÉMY

- *užitočnosť NS (učenie a generalizácia)*
- *vlastnosti a schopnosti NS*
- *architektúry NS*
- *návrh neurónovej siete*
- *pravidlá pre porovnanie podobnosti vstupov neurónovej siete*
- **UČENIE NEURÓNOVÝCH SIETÍ**
  - *algoritmy učenia (learning algorithms)*
    - *učenie korigujúce chybu (error-correction learning)*
    - *Hebbovo učenie (Hebbian learning)*
    - *súťažné učenie (competitive learning)*

- *paradigmy učenia*
  - *učenie s učiteľom (supervised learning)*
  - *známkované učenie (reinforcement learning)*
  - *učenie bez učiteľa, samoorganizácia (unsupervised learning, self-organization)*
- *vhodné úlohy pre učenie neurónových sietí*
  - *aproximácia*
  - *asociácia*
  - *klasifikácia vzorov*
  - *predikcia*
  - *riadenie*
  - *separácia vzorov*
  - *...*

# PREDNÁŠKA

- užitočnosť NS

- učenie
- generalizácia

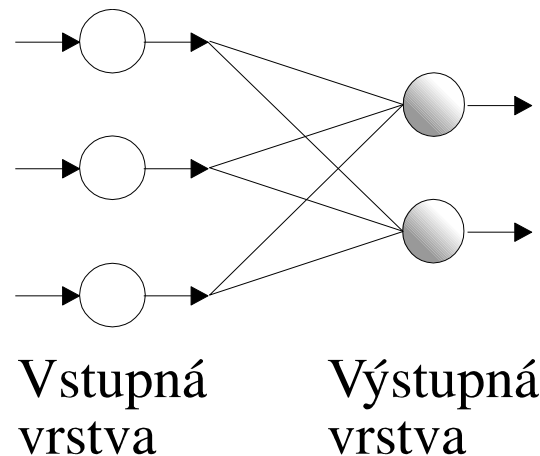
- vlastnosti a schopnosti NS

- nelinearita
  - nelinearita je distribuovaná v sieti
- vstupno-výstupné mapovanie (input-output mapping)
  - sieť si nastavuje synaptické váhy tak, aby minimalizovala rozdiel medzi žiadanou a aktuálnou odpoveďou

- adaptívnosť
  - adaptovanie synaptických váh na zmeny okolitého prostredia
  - adaptovanie možné aj v reálnom čase
- odolnosť voči chybám
  - postupná degradácia pri nepriaznivých podmienkach činnosti
  - informácie sú v sieti distribuované - až veľmi veľké poškodenie vedie k úplnej degradácii odpovedí siete
- VLSI implementácia
- analógia s neurobiológiou
- ...

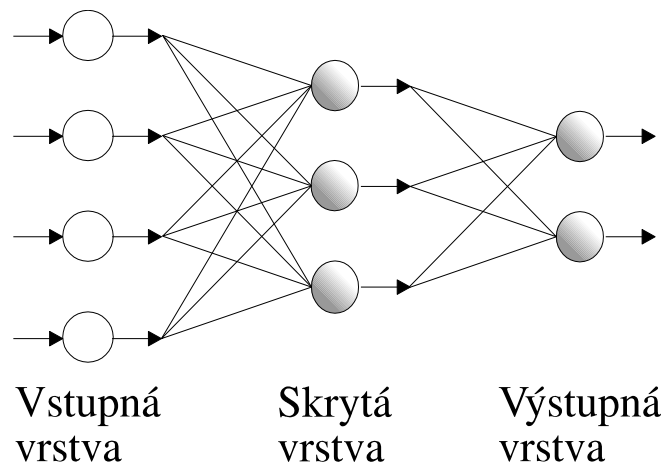
- architektúry NS

- jednovrstvové dopredné siete (single-layer feedforward networks)
  - sieť konfigurácie  $p - q$ , kde  $p$  je počet neurónov vo vstupnej vrstve a  $q$  počet neurónov vo výstupnej vrstve



*dopredná sieť konfigurácie 3-2*

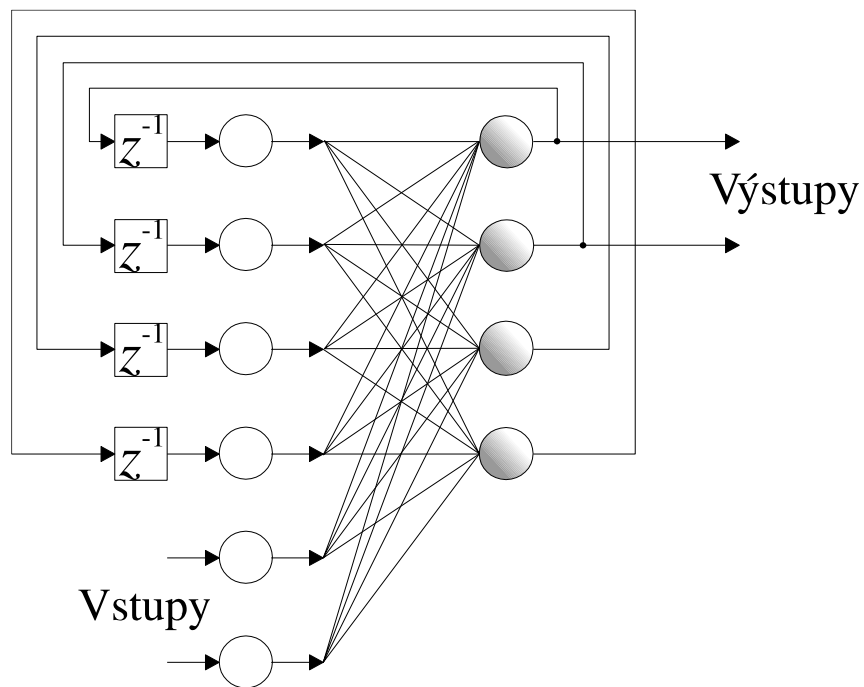
- viacvrstvové dopredné siete (multi-layer feedforward networks)
  - jednu alebo viac *skrytých vrstiev* (“hidden layer”)
  - neuróny v nich sa nazývajú skryté neuróny
  - sieť konfigurácie  $p - h_1 - h_2 - q$
  - sieť nazývame *úplne prepojenou*, ak každý neurón v každej vrstve je prepojený s každým neurónom v predchádzajúcej vrstve



*úplne prepojená dopredná sieť konfigurácie 4-3-2*

- receptívne a projektívne polia

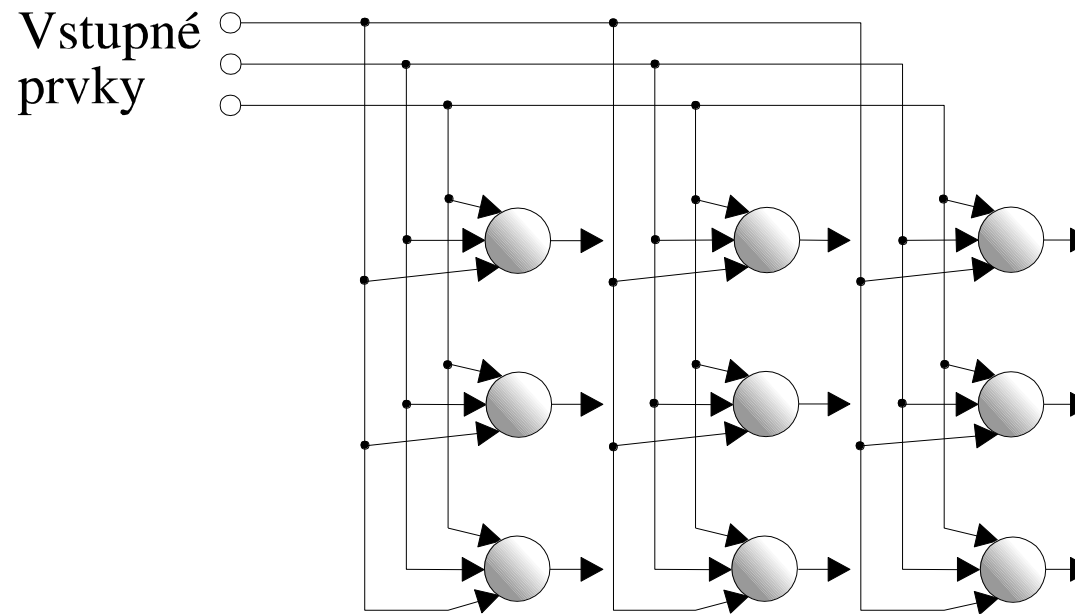
- rekurentné siete (recurrent networks)
  - majú aspoň jednu spätnú väzbu
  - spätnoväzbové slučky využívajú oneskorenia: značenie  $z^{-1}$



*rekurentná sieť so skrytými neurónmi*



- mriežkové štruktúry (lattice structures)
  - 1D, 2D alebo viacrozmerné pole neurónov so zodpovedajúcimi vstupnými prvkami
  - každý vstupný prvok je spojený s každým neurónom v mriežke



*dvojrozmerná mriežka s 3x3 neurónmi*

- návrh neurónovej siete

- výber vhodnej architektúry
  - tréovanie siete na podmnožinu dostupných tréovacích údajov pomocou vhodného algoritmu (učenie)
  - testovanie správania sa siete pre dáta, ktoré sieť pri učení nemala k dispozícii (generalizácia)
- zásadný rozdiel medzi návrhom neurónovej siete a klasickým prístupom k spracovaniu informácií
    - klasický prístup - vytvorenie matematického modelu a jeho overenie na reálnych dátach
    - návrh neurónovej siete je založený priamo na reálnych dátach, ktoré “hovorí samé za seba”

- pravidlá pre porovnanie podobnosti vstupov neurónovej siete

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{iN}]^T$$

skalárny súčin

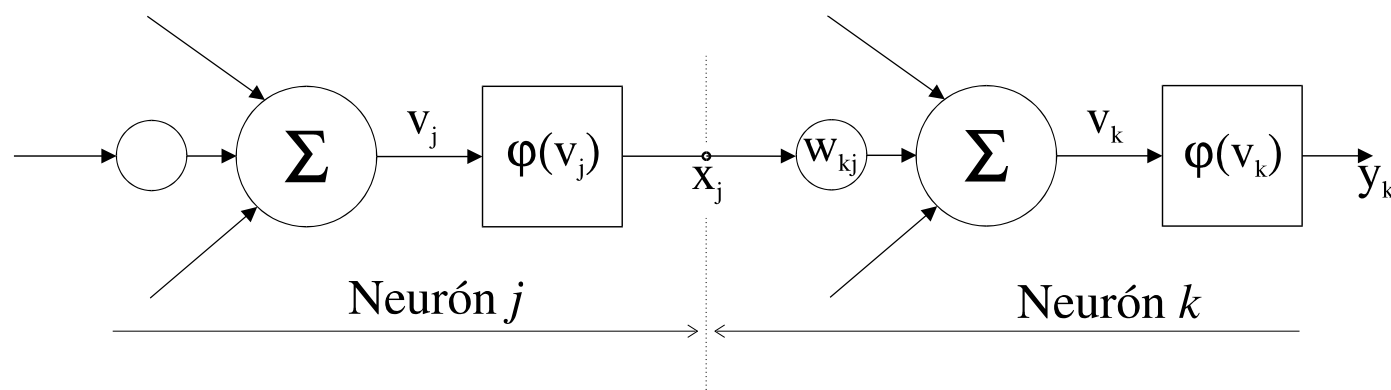
$$\mathbf{x}_i^T \mathbf{x}_j = \sum_{n=1}^N x_{in} x_{jn}$$

euklidovská vzdialenosť

$$\|\mathbf{x}_i - \mathbf{x}_j\| = \left[ \sum_{n=1}^N (x_{in} - x_{jn})^2 \right]^{1/2}$$

## UČENIE NEURÓNOVÝCH SIETÍ

- učenie je proces, ktorým sa adaptujú voľné parametre neurónovej siete pomocou stimulácie prostredím, v ktorom neurónová sieť pracuje



algoritmus učenia:

$$w_{kj}(n+1) = w_{kj}(n) + \Delta w_{kj}(n)$$

Iteratívne algoritmy pre učenie neurónov a z nich vytvorených sietí rešpektujú **princíp minimálneho** narušenia (“minimal disturbance principle”): Adaptuj tak, aby sa zredukovala výstupná chyba pre súčasnú trénovaciu vzorku a aby sa minimálne narušili už naučené vzorky.

- algoritmy učenia (learning algorithms)

- učenie korigujúce chybu (error-correction learning)

- $d_k(n)$  - požadovaná odpoveď pre neurón  $k$  v čase  $n$
    - $y_k(n)$  - aktuálna odpoveď - odpoveď na vstupný vektor  $\mathbf{x}(n)$
    - vstupný vektor  $\mathbf{x}(n)$  a žiadaná odpoveď  $d_k(n)$  pre neurón  $k$  tvoria *príklad* (vzorku) pre sieť v čase  $n$

chybový signál 
$$e_k(n) = d_k(n) - y_k(n)$$

stredná kvadratická chyba 
$$J = E \left[ \frac{1}{2} \sum_k e_k^2(n) \right]$$

okamžitá kvadratická chyba 
$$\varepsilon(n) = \frac{1}{2} \sum_k e_k^2(n)$$

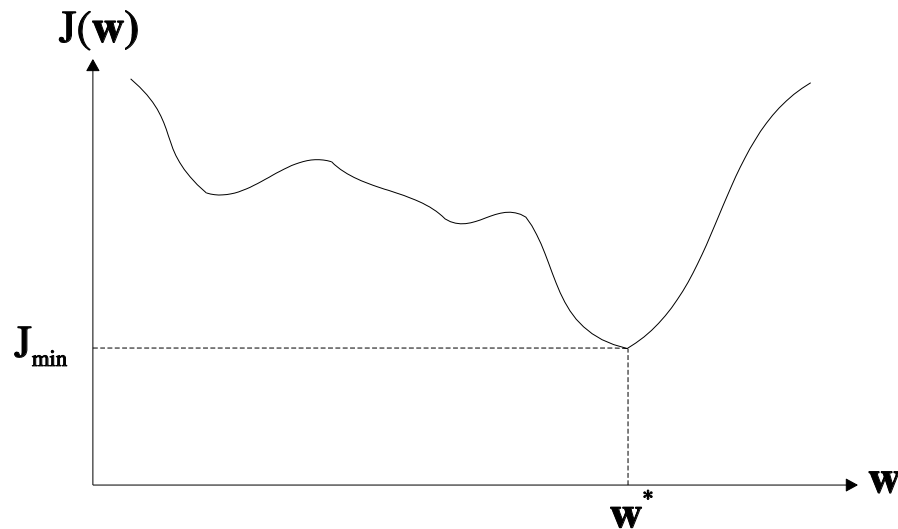
- *pravidlo korekcie chyby* (“error-correction learning rule”), niekedy nazývané aj *delta pravidlo*

$$\Delta w_{kj}(n) = \eta e_k(n) x_j(n)$$

zmena váhy = (parameter rýchľ. učenia) x (chyba) x (vstup)

- metóda gradientového zostupu (“method of gradient descent”)
- parameter rýchlosti učenia
  - stabilita procesu učenia
    - ak je hodnota malá, učenie je bez problémov, avšak konvergencia je veľmi pomalá
    - veľká hodnota znamená rýchle učenie, je však možnosť, že učenie bude divergovať a systém sa stane nestabilným

- chybový povrch (error surface)
  - graf kritériálnej funkcie  $J$  voči synaptickým váham siete



- 1) Neurónová sieť obsahuje iba lineárne prvky. Chybový povrch je vtedy kvadratickou funkciou váh siete a obsahuje **jediné** minimum.
- 2) Neurónová sieť obsahuje nelineárne prvky. V tomto prípade má chybový povrch **globálne minimum** (prípadne viacnásobné globálne minimum) a **aj lokálne minimá**.

ilustrácia chybových povrchov pre hlboké siete: [Tom Goldstein: "What do neural loss surfaces look like?" - YouTube](#)

- Hebbovo učenie (Hebbian learning)

- Hebbovo pravidlo:

1. Ak dva neuróny na oboch stranách synapsy (spojenia) sú aktivované súčasne (synchronne), potom sila synapsy sa zväčší.
2. Ak dva neuróny na oboch stranách synapsy sú aktivované asynchronne, potom je táto synapsa zoslabená alebo eliminovaná.

- *Hebbovská synapsa* zvyšuje silu spojenia s kladne korelovanou presynaptickou a postsynaptickou aktivitou a znižuje silu spojenia ak tieto aktivity sú nekorelované alebo záporne korelované.
- *Antihebbovská synapsa* zoslabuje kladne korelované presynaptické a postsynaptické aktivity a zosilňuje záporne korelované aktivity.
- *Nehebbovská synapsa* takéto mechanizmy nezahŕňa.



$$\Delta w_{kj}(n) = F(y_k(n), x_j(n))$$

$$\Delta w_{kj}(n) = \eta y_k(n) x_j(n)$$

zmena váhy = (parameter rýchľ. učenia) x (výstup) x (vstup)

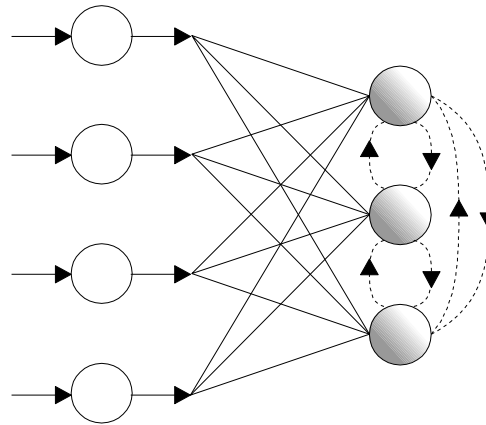
$$\Delta w_{kj}(n) = \eta y_k(n) x_j(n) - \alpha y_k(n) w_{kj}(n)$$

$$\Delta w_{kj}(n) = \eta \text{cov}[y_k(n), x_j(n)] = \eta E[(y_k(n) - \bar{y}_k)(x_j(n) - \bar{x}_j)]$$

$$\Delta w_{kj}(n) = \eta \{E[y_k(n) x_j(n)] - \bar{y}_k \bar{x}_j\}$$

- súťažné učenie (competitive learning)

- jedna vrstva výstupných neurónov - úplne prepojená so vstupnými prvkami
- laterálne spojenia - slúžia pre laterálnu inhibíciu (každý neurón snaží “utlmiť” neuróny, s ktorými je spojený)
- dopredné spojenia – excitačné



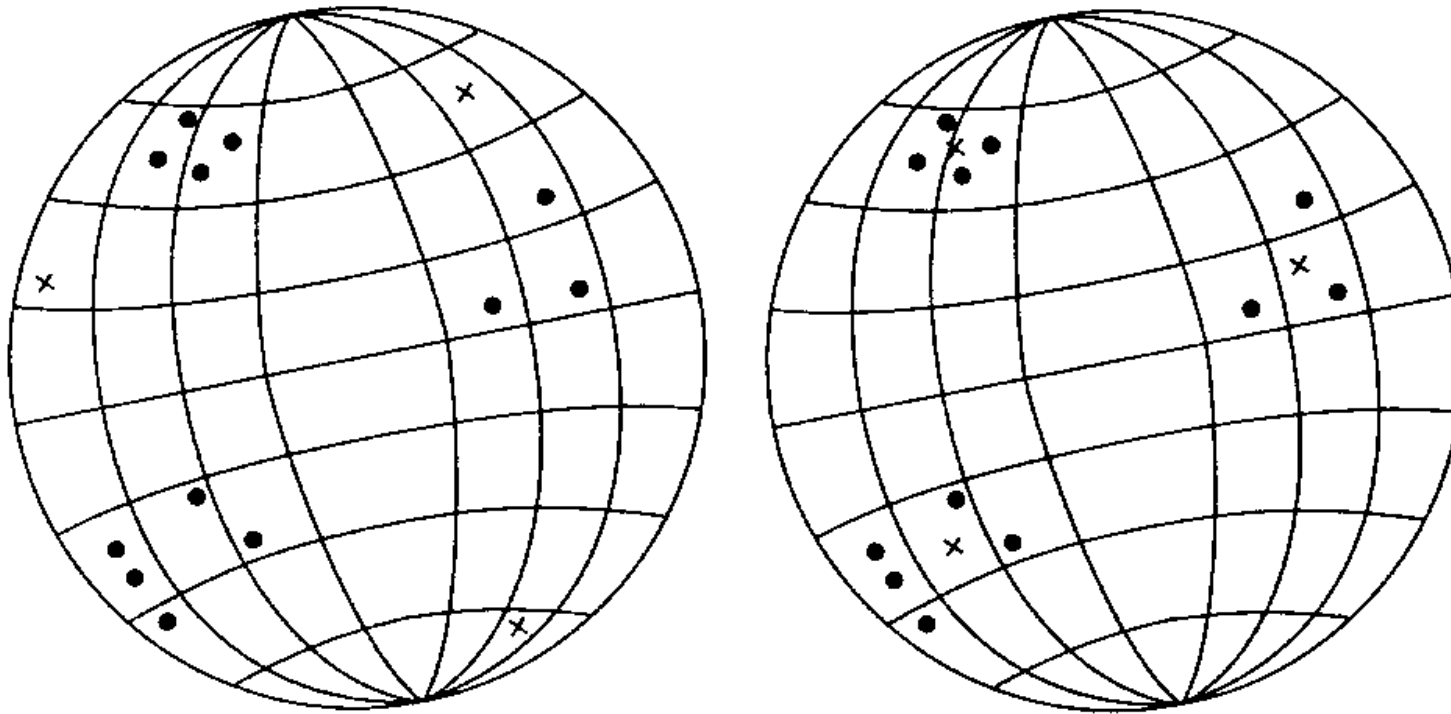
*sieť so súťažným učením s doprednými (excitačnými) a laterálnymi (inhibičnými) spojeniami*

- výstupné neuróny siete vzájomne súťažia s cieľom stať sa aktívnym neurónom
- sieť s Hebbovým učením - viaceré výstupné neuróny môžu byť aktívne súčasne
- súťažné učenie - aktívny iba jeden neurón
  - neurón typu “vítaz berie všetko” (“winner-takes-all neuron”), vítaz (“winner”)

zmena váhy = (parameter rýchľ. učenia) x [ (vstup) - (váha od tohto vstupu) ]

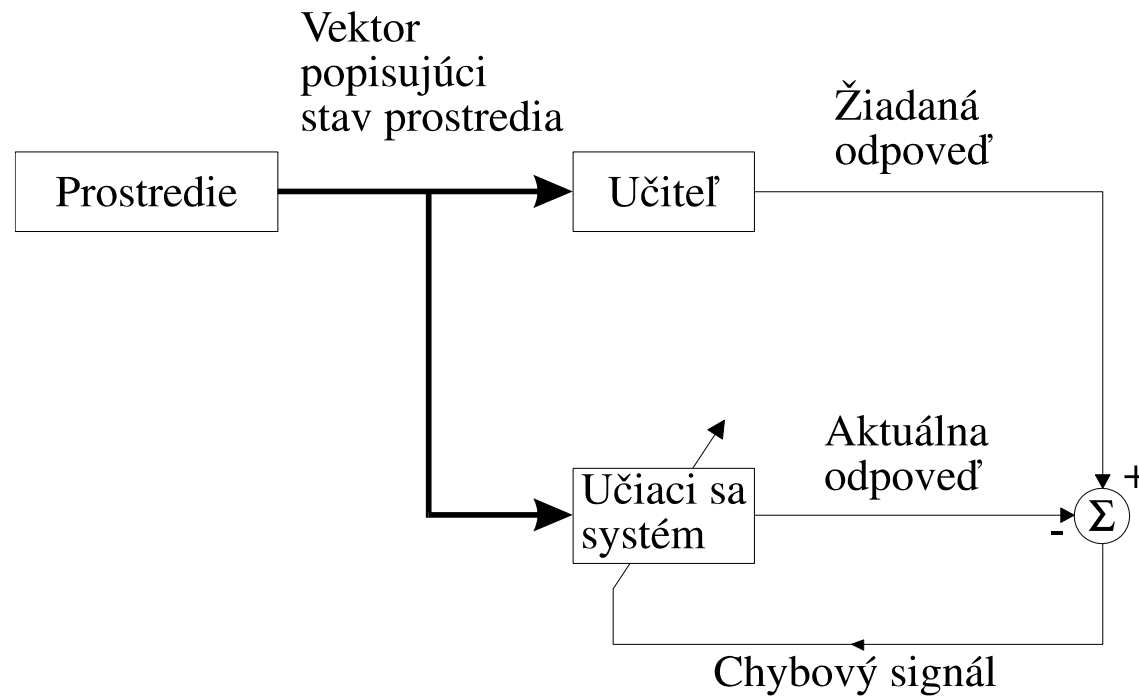
$$\Delta w_{kj}(n) = \begin{cases} \eta [x_j(n) - w_{kj}(n)], & \text{ak neurón } k \text{ vyhrá súťaž} \\ 0, & \text{ak neurón } k \text{ nevyhrá súťaž} \end{cases}$$

- účinok -posun váhového vektora  $\mathbf{w}_k$  víťaza k bližšie smerom k vstupnej vzorke  $\mathbf{x}$



*Geometrická interpretácia procesu so súťažným učením - počiatočný a konečný stav siete*

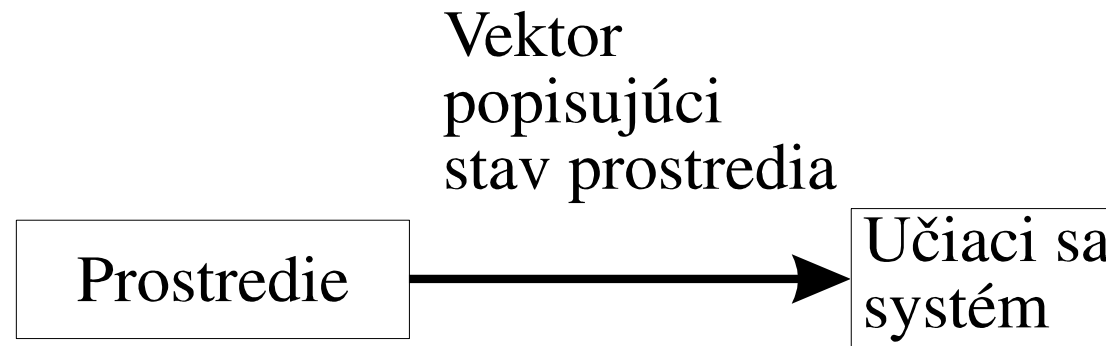
- paradigmy učenia
  - učenie s učiteľom (supervised learning)



napr. LMS (least-mean square) algoritmus a jeho zovšeobecnenie - algoritmus spätného šírenia ("backpropagation algorithm")

- známkované učenie (reinforcement learning)
  - nie požadované odpovede, ale známky (hodnotenie odpovedí)

- učenie bez učiteľa, samoorganizácia (unsupervised learning, self-organization)



- žiadny učiteľ, žiadne ohodnotenie
- sieti sa poskytne iba jediné - meradlo kvality pre reprezentáciu, ktorú sa má sieť naučiť

- vhodné úlohy pre učenie neurónových sietí

- aproximácia

nelineárne vstupno-výstupné mapovanie  $d = g(\mathbf{x})$  ( $g$  nepoznáme)  
sú zadané príklady vo forme vstupno-výstupných párov  $(\mathbf{x}_1, d_1), (\mathbf{x}_2, d_2), \dots, (\mathbf{x}_N, d_N)$   
učenie s učiteľom

- asociácia

autoasociácia - samoorganizované učenie  
heteroasociácia - učenie s učiteľom



- klasifikácia vzorov

daný pevný počet tried (kategórií), do ktorých majú byť vstupy klasifikované  
učenie s učiteľom

keď nie je znalosť o triedach – zhľukovanie - samoorganizácia

- predikcia

$M$  predchádzajúcich vzoriek  $x(n-1), x(n-2), \dots, x(n-M)$

nasledujúca vzorka  $x(n) = ?$

chyba predikcie

$$e(n) = x(n) - \hat{x}(n|n-1, \dots, n-M)$$

učenie s učiteľom, ale nasledujúca vzorka je jasná z časového radu - samoorganizácia

- riadenie
- separácia vzorov
- ...

- Dobrá ilustrácia činnosti umelého neurónu pri klasifikácii dát
  - <http://playground.tensorflow.org>

