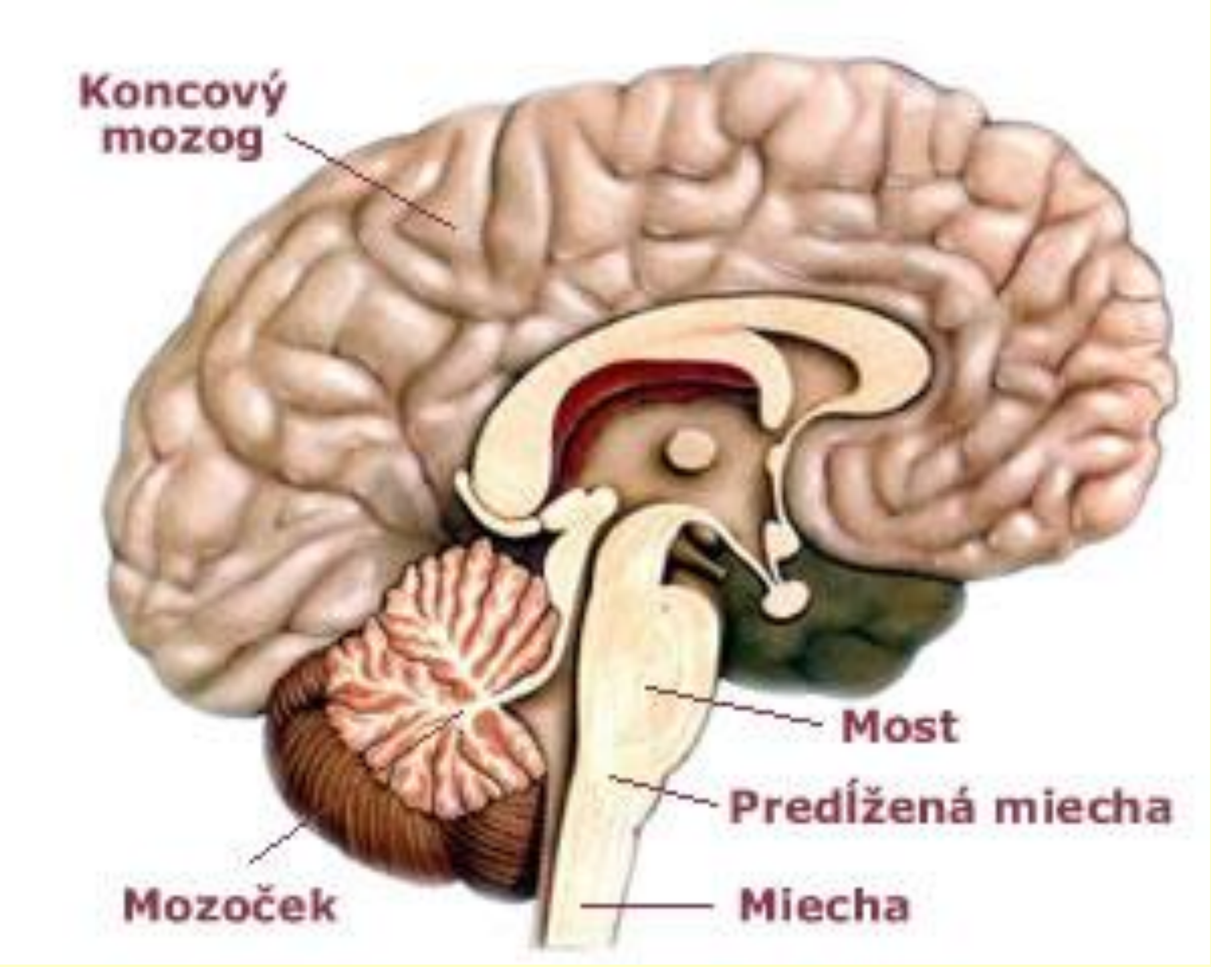
Úvod do neurónových sietí

Matematický model neurónu a neurónovej siete. Postavenie neurónových sietí v informatike.

Prahové jednotky.

Perceptróny. Lineárne separovateľné objekty, adaptačný proces (učenie), konvergencia perceptrónu, perceptróny vyššieho rádu

Čo vieme o mozgu?



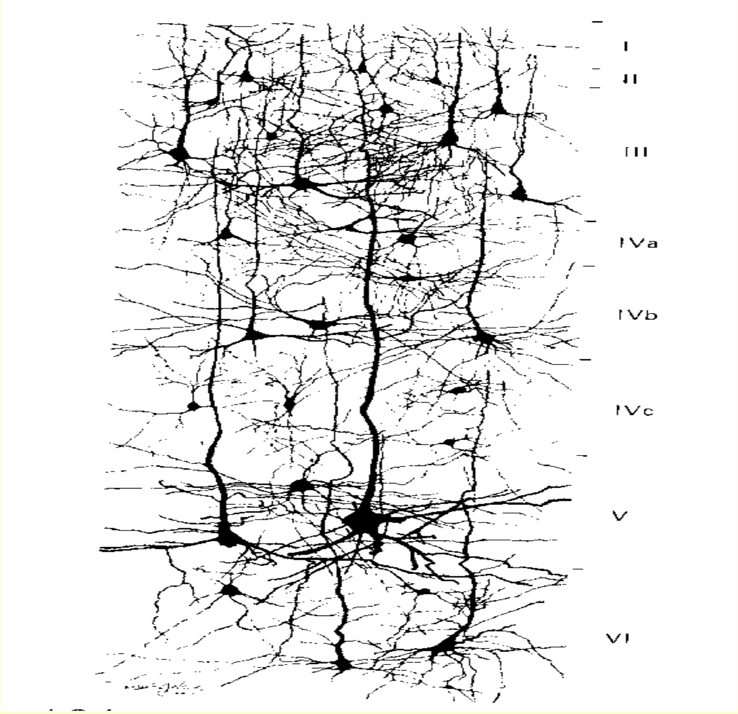
Mýty o mozgu

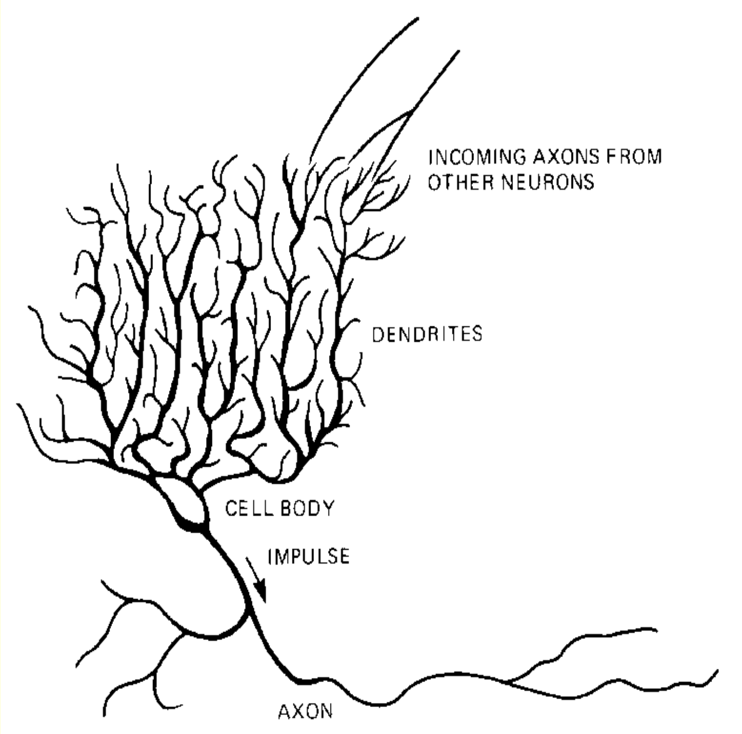
- Aby sme mohli účinne fungovať potrebujeme celý mozog.
- Niektorí ľudia, ktorým bola odobratá v detstve kvôli chorobe jedna hemisféra, môžu v dospelosti fungovať primerane dobre.
- Moderný človek má väčší mozog než neandertálci.
- Mozog neandertálca bol pravdepodobne o niečo väčší než náš.
- U dospelých sa nevytvárajú nové neuróny.
- Pomerne nedávne výskumy poukazujú na rast nových neurónov v niektorých častiach mozgu, hlavne v hypokampe.
- V dospelosti stratíme každý deň zhruba 100 tisíc neurónov.
- Pravdepodobne desaťkrát menej.

Princíp fungovania mozgu

- V 20. storočí. V roku 1926 anglický fyzik W. Grey zistil, že existujú určité frekvenčné hladiny mozgu. Nazval ich **alfa, beta, delta a theta**.
- Každá hladina vedomia pracuje na určitej frekvencii, respektíve rozpätí frekvencií. Ľudský mozog teda funguje na prenose malých elektrických výbojov v rámci neurotransmiterov.
- Mozog je vždy naladený na jednu z týchto hladín vedomia, a to podľa toho, ako sa cítime alebo aké zmyslové orgány používame, prípadne nepoužívame.
- Mozgová frekvencia je merateľná v Hertzoch (Hz) prístrojom EEG (elektroencefalograf).

Ako vyzerá biologický neurón?





Prvý obrázok znázorňuje neuróny v mozgu a ich veľkú vzájomnú prepojenosť.

Druhý obrázok znázorňuje neurón (nerve cell) zložený z viacerých vstupných spojení (dendrites) a jedného silnejšieho výstupného spojenia (axon).

Fenomén neurónových sietí (NS)

- sú inšpirované biologickými neurónovými sieťami
- využívajú distribuované, paralelné spracovanie informácie pri vykonaní výpočtov
- znalosti sú ukladané predovšetkým prostredníctvom sily väzieb medzi jednotlivými neurónmi
- učenie je základná a podstatná vlastnosť neurónových sietí

O učení

- V mozgovej kôre a vôbec v celom mozgu je úplne geneticky určené, ktoré neuróny a akým typom synapsií budú komunikovať.
- Koľko synapsií bude medzi danými neurónmi, je tiež podstatne geneticky určené.
- V priebehu celého života sa v dôsledku individuálnej skúsenosti (učenia) menia váhy jednotlivých synapsií.
- Otázka: Akými pravidlami sa tieto úpravy váh riadia?

História umelých neurónových sietí

- Warren Mc Culloch, Walter Pitts, 1943
- najjednoduchšie typy NS ... ľubovoľná aritmetická alebo logická funkcia
- Donald Hebb: The Organization of Behavior, 1949
- učiace pravidlo pre synapsie neurónov
- prvý neuropočítač Smark, 1951, Marwin Minsky
- Frank Rosenblatt, 1957... perceptrón, učiaci algoritmus - Principles of Neurodynamics

História

- 1957, 1958 Mark I - Perceptron
- ADALINE
- Nils Nilsson: Learning Machines, 1965 – kritický postoj
- 1967-1982 - ojedinelý výskum, NS zaznávané
- 80. roky - rozvoj
- 1982, 84 John Hopfield – vyznačne sa zaslúžil o pokračovanie vo výskume NS

Umelé neurónové siete sú vlastne

- **paralelné výpočtové modely -** husto poprepájané adaptívne jednotky (procesory)
- **majú adaptívnu povahu – "učenie z príkladov" -** čo je veľmi žiadané v aplikáciách

Predpokladáme

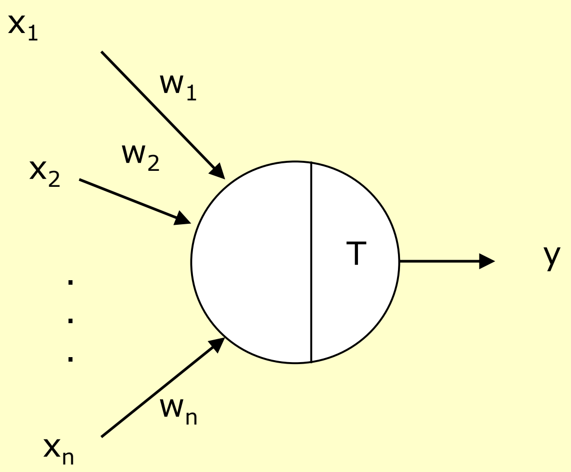
- sú k dispozícii tréningové údaje
- je neúplné porozumenie riešeného problému (napríklad nevieme tvar funkcie, ktorú potrebujeme nájsť)

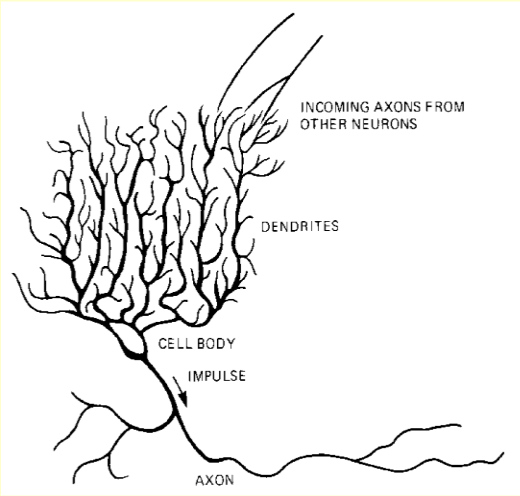
Výpočtové modely pre triedy problémov:

- klasifikácia
- syntéza a rozpoznávanie reči
- aproximácia funkcií
- kompresia údajov
- vyhľadávanie klastrov
- predikcia
- kombinatorické optimalizačné úlohy
- modelovanie nelineárnych systémov
- riadenie

Prahové jednotky (Threshold Gates)

- **Lineárne prahové jednotky - LTG**





Čo počíta LTG

- Základná funkcia - rozlišovať body (vektory) patriace do dvoch tried.
- Prenosová funkcia LTG je daná aj analyticky

$$y = \begin{cases} 1 & \text{ak} \quad \mathbf{w}^T\mathbf{x} = \sum_{i=1}^{n} w_i x_i \geq T \\ 0 & \text{inak} \end{cases}$$

kde

$\mathbf{x}=(x_1,x_2,\ldots,x_n)^T$	sú vstupné hodnoty
$\mathbf{w}=(w_1,w_2,\ldots,w_n)$	hodnoty váh
$\boldsymbol{T}$	je prahová hodnota.

Čo počíta LTG

- Vektor **x** je **n** - rozmerný s binárnymi alebo reálnymi zložkami
- $\mathbf{x} \in (0,1)^n$ alebo $\mathbf{x} \in R^n$
- váhy $\mathbf{w} \in R^n$
- prah je z R

- LTG počíta jedno z nasledujúcich zobrazení:
- $\mathbf{x} \in (0,1)^n \Rightarrow y: \{0,1\}^n \to \{0,1\}$ - Booleovské funkcie
- $\mathbf{x} \in R^n \quad \Rightarrow y: R^n \to \{0,1\}$ - Klasifikácia do dvoch tried

Príklad realizácie booleovskej funkcie pomocou LTG

- $y(x1, x2, x3) = \underline{x1}\ x2 + x2\ \underline{x3}$.
- pomocou LTG, pričom podčiarknuté sú negácie.

- **Problém** : Nie je možné skonštruovať LTG pre booleovskú funkciu XOR - $y(x1, x2)$, t.j. funkciu

$y(0,0) = y(1,1) = 0, \quad y(0,1) = y(1,0) = 1$

Prahová funkcia

- Booleovská funkcia, ktorá je realizovateľná pomocou jednej LTG sa nazýva *prahová funkcia*.

- Prahová funkcia je vlastne *lineárne separovateľná funkcia*, t. j. funkcia so vstupmi patriacimi do dvoch rôznych tried oddelených nadrovinou (takými, že vstupy jednej triedy môžu byť geometricky separované od vstupov druhej triedy nadrovinou).

Porovnanie počtov prahových funkcií a. počtov všetkých možných Booleovských funkcií pre vybrané hodnoty n.

n	B_n	Total number of BF
1	4	4
2	14	16
3	104	256
4	1882	65536
5	94572	$4.3*10^{9}$
6	15028134	$1.8*10^{19}$

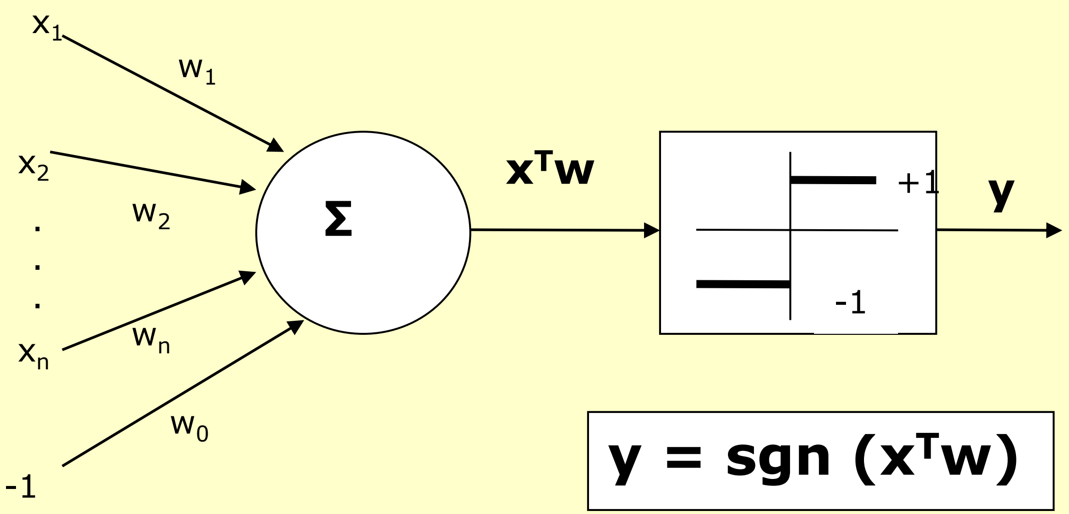
Polynomiálne prahové jednotky

Y=1, ak

$$\sum_{i_1=1}^{n} w_{i_1} x_{i_1} + \sum_{i_1=1}^{n} \sum_{i_2=i_1}^{n} w_{i_i i_2} x_{i_1} x_{i_2} + \ldots + \sum_{i_1=1}^{n} \ldots \sum_{i_r=i_{r-1}}^{n} w_{i_1 i_2 \ldots i_r} x_{i_1} x_{i_2} \ldots x_{i_r} \geq T$$

Y=0, inak

Perceptróny, učiace pravidlá



Učiace pravidlá pre perceptróny

- Predpokladajme, že trénujeme perceptrón pomocou **tréningovej množiny** – dvojice

$\{ (\mathbf{x}_1, d_1), (\mathbf{x}_2, d_2), ..., (\mathbf{x}_m, d_m) \}$,

- kde $\mathbf{x_k} \in R^{n+1}$ je k-tý vstupný vektor a $d_k \in \{-1,+1\}$, k=1, 2,..., m je predpokladaná trieda (cieľ) pre k - tý vstupný vektor (poradie dvojíc - náhodné)

Učiace pravidlá pre perceptróny

Správna klasifikácia

- Cieľ - navrhnúť perceptrón, ktorý pre každý vstupný vektor $\mathbf{x_k}$ tréningovej množiny dá výstup y_k, ktorý sa zhoduje s cieľom d_k. Teda $y_k = \mathrm{sgn}(\mathbf{w}^T\mathbf{x}_k) = d_k$, k=1, 2,..., m.

Riešenie $\mathbf{w}^*$ - **učiace pravidlo (Rosenblatt, 1962)**:

- $\mathbf{W_0}$ = ľubovoľné
- $\mathbf{W}_{k+1} = \mathbf{W}_k + ro\,(d_{k+1} - y_{k+1})\mathbf{x}_{k+1}$, k=1,2,...

  kde ro je kladná konštanta - **učiaci pomer** z otvoreného intervalu (0, 1).

Príklad

- Máme dané trénujúce príklady

m	x^m	d^m
1	(2,6)	1
2	(-1,-4)	-1
3	(2,2)	1
4	(1,-3)	-1
5	(-3,-1)	1

Je potrebné nájsť perceptrón, ktorý bude správne klasifikovať uvedené body v rovine.

Príklad

Krok\Prem.	m	k	ρ	x^m	d^m	y^m	$d^m=x^m$	nezhoda	w	m< M	nezhoda = false
1	1	1	0,5					FALSE	(10, -1, -2)		
2				$x^1=(2,6)$	$d^1=1$	y^1=sgn(10.(-1)+(-1).2+(-2).6)= -1					
3							-	TRUE	(10, -1, -2)+1/2(1-(-1)).(-1, 2, 6) = (9, 1, 4)		
4	2									+	
5				$x^2=(-1,-4)$	$d^2=-1$	y^2=sgn(9.(-1)+ 1.(-1)+ 4.(-4))=-1					
6							+		(9, 1, 4)		
7	3									+	
8				$x^3=(2,2)$	$d^3=1$	y^3=sgn(9.(-1)+ 1.2 + 4.2)=1					
9							+		(9, 1, 4)		
10	4									+	
11				$x^4=(1,-3)$	$d^4=-1$	y^4=sgn(9.(-1)+ 1.1+ 4.(-3))= -1					
12							+		(9, 1, 4)		
13	5									+	
14				$x^5=(-3,-1)$	$d^5=1$	y^5=sgn(9.(-1)+ 1.(-3)+ 4.(-1))= -1					
15							-	TRUE	(9, 1, 4)+1/2(1-(-1)).(-1, -3, -1) = (8, -2, 3)		
16	6									+	
17				$x^6=(-1,1)$	$d^6=1$	y^6=sgn(8.(-1)+ (-2).(-1)+ 3.1)= -1					
18							-	TRUE	(8, -2, 4)+1/2(1-(-1)).(-1, -1, 1) = (7, -3, 4)		
19	7									+	
20				$x^7=(6,1)$	$d^7=-1$	y^7=sgn(7.(-1)+ (-3).6+ 4.1)= -1					
21							+		(7, -3, 4)		
22										-	
23	1	2									-

m	x^m	d^m
1	(2,6)	1
2	(-1,-4)	-1
3	(2,2)	1
4	(1,-3)	-1
5	(-3,-1)	1

- tréningová množina obsahuje príklady zapísané v Tab.
- počiatočné nastavenie váh je = (-1, -2) ; hodnota prahu =10;
- učiaci pomer = 0,5

Poznámky

- ak riešenie existuje, potom je dosiahnuté po konečnom počte iterácií;
- vo všeobecnosti táto metóda poskytuje nekonečný počet riešení;
- Rosenblattovo učiace pravidlo konverguje k jednému z nich;
- riešenie je citlivé na hodnotu učiaceho pomeru ro, ktorá koriguje prezentáciu dvojice;
- počet korekcií k0 závisí od výberu počiatočných hodnôt vektora váh $\mathbf{W}_0$;
- ak riešenie neexistuje, perceptrón to neodhalí
- vektor riešenia existuje, vtedy a len vtedy daná tréningová množina je lineárne separovateľná

Poznámky

Veta: [Novikoff, 1962; Ridgway, 1962; Nilsson,1965]}

Ak tréningová množina je lineárne separovateľná, potom učiace pravidlo pre perceptróny konverguje k riešeniu v konečnom počte iterácií.