

Modely neurónových sietí, ich schopnosti a aplikácie

Gabriela Andrejková

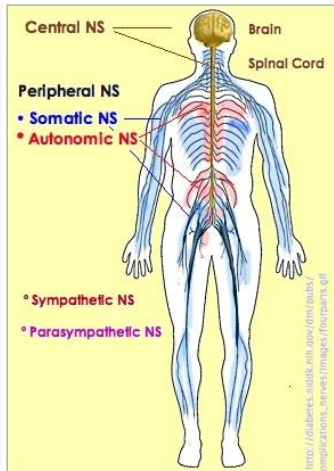
Ústav informatiky, Prírodovedecká fakulta, UPJŠ, Košice



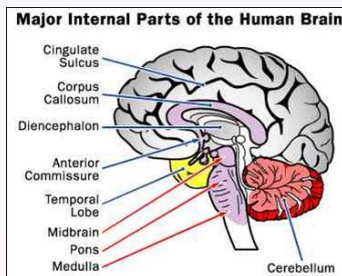
Obsah prednášky

1. Prečo neurónové siete
2. Čo dokáže jeden neurón
3. Vrstvové siete a aproximácia funkcií
4. Hopfieldove siete, riešenie optimalizačných úloh
5. Polynomiálne neurónové siete

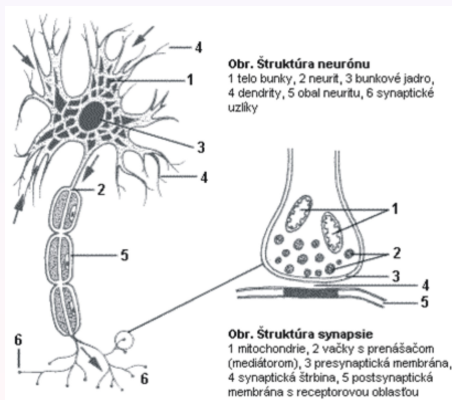
Pripomeňme si naše neuróny a naše neurónové siete



Pripomeňme si naše neuróny a naše neurónové siete



Pripomeňme si naše neuróny a naše neurónové siete

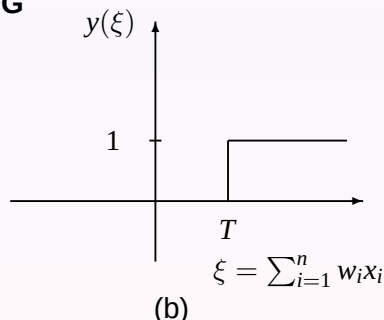
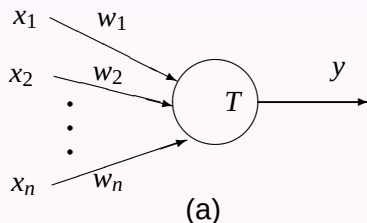


Fenomén neurónových sietí

- ▶ sú inšpirované biologickými neurónovými sietami
- ▶ 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

Prahové jednotky (Threshold Gates)

Lineárne prahové jednotky - LTG



(a) Grafické znázornenie lineárnej prahovej jednotky a
(b) jej prenosovej funkcie.

Prahové jednotky (Threshold Gates)

Lineárne prahové jednotky - LTG

Základná funkcia - rozlíšovať body (vektory) patriace do dvoch tried.

Prenosová funkcia LTG je daná analyticky

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

$\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ sú vstupné hodnoty,

$\mathbf{w} = [w_1, w_2, \dots, w_n]$ hodnoty váh,

T je prahová hodnota.

Lineárne prahové jednotky - LTG

Vektor $\mathbf{x}$ je n -rozmerný s binárnymi alebo reálnymi zložkami

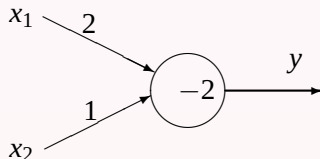
- ▶ $\mathbf{x} \in \{0, 1\}^n$, alebo $\mathbf{x} \in \mathcal{R}^n$,
- ▶ váhy $\mathbf{w} \in \mathcal{R}^n$.

LTG počíta jedno z nasledujúcich zobrazení:

- ▶ $\mathbf{x} \in \{0, 1\}^n \Rightarrow y : \{0, 1\}^n \rightarrow \{0, 1\}$,
- ▶ $\mathbf{x} \in \mathcal{R}^n \Rightarrow y : \mathcal{R}^n \rightarrow \{0, 1\}$,

Príklad LPG pre booleovskú funkciu 2 premenných

x_1	x_2	y
0	0	0
0	1	0
1	0	1
1	1	1



Problém: skonštruovať LTG pre booleovskú funkciu XOR

Definícia XOR:

$$y(x_1, x_2), y(0, 0) = y(1, 1) = 0, y(0, 1) = y(1, 0) = 1.$$

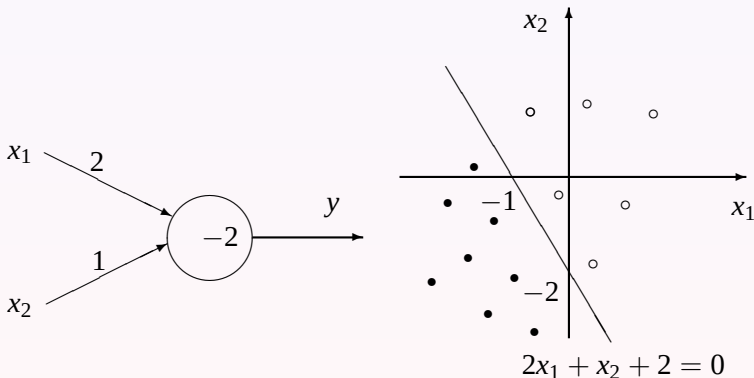
- ▶ Booleovská funkcia, ktorá je realizovateľná pomocou jednej LTG sa nazýva *prahová funkcia*.
- ▶ Prahová funkcia je **lineárne separovateľná funkcia**, t.j. funkcia so vstupmi patriacimi do dvoch rôznych tried, takými že vstupy jednej triedy môžu byť geometricky separované od vstupov druhej triedy nadrovinou.

Porovnanie počtov prahových funkcií vs. počtov všetkých možných boolovských funkcií pre vybrané hodnoty n .

n	B_n	Počet BF 2^{2^n}
1	4	4
2	14	16
3	104	256
4	1882	65536
5	94572	4.3×10^9
6	15028134	1.8×10^{19}

$$\lim_{n \rightarrow \infty} \frac{B_n}{2^{2^n}} = 0$$

Príklad separujúcej funkcie pomocou LTG - 2 triedy (reálne vstupy)



- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

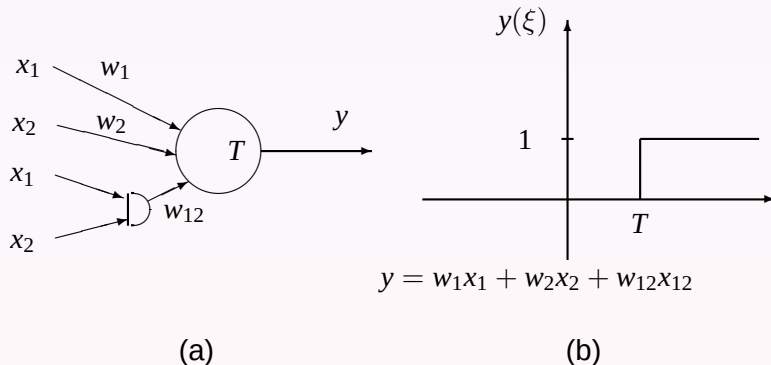
KVADRATICKÉ prahové jednotky

$$y = \begin{cases} 1 & \text{if } \sum_{i=1}^n (w_i x_i + \sum_{j=i}^n w_{ij} x_i x_j) \geq T, \\ 0 & \text{inak} \end{cases} \quad (2)$$

kde $\mathbf{x} = [x_1 x_2 \dots x_n]^T$ sú vstupy a

$\mathbf{w} = [w_1, \dots, w_n, w_{11} \dots w_{nn}]$ sú hodnoty váh, T je prahová hodnota.

(a) Kvadratická prahová jednotka a
(b) jej prenosová funkcia.



XOR : $w_1 = 0.3, w_2 = 0.3, w_{12} = -0.5, T = 0.2$

POLYNOMIÁLNE prahové jednotky - PTG

$$y = \begin{cases} 1 & \text{ak } \sum_{i_1=1}^n w_{i_1} x_{i_1} \\ & + \sum_{i_1=1}^n \sum_{i_2=i_1}^n w_{i_1 i_2} x_{i_1} x_{i_2} + \dots \\ & + \sum_{i_1=1}^n \dots \sum_{i_r=i_{r-1}}^n w_{i_1 i_2 \dots i_r} x_{i_1} x_{i_2} \dots x_{i_r} \geq T, \\ 0 & \text{inak} \end{cases} \quad (3)$$

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

Výpočtové schopnosti PTG pre booleovské funkcie

Theorem (Nilsson, 1965; Krishnan, 1966)

Ľubovoľná booleovská funkcia n premenných môže byť realizovaná pomocou polynomiálnej prahovej jednotky rádu r , $r \leq n$.

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

Výpočtové schopnosti PTG pre aproximácie funkcií

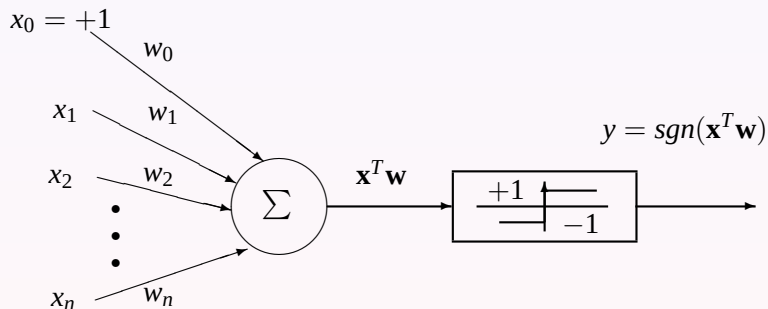
Theorem (Weierstrass)

Nech g je spojitá reálna funkcia, definovaná na uzavretom intervale $[a, b]$. Potom pre ľubovoľné $\varepsilon > 0$, existuje polynóm y (ktorý môže závisieť od ε) s reálnymi koeficientami taký, že pre každé $x \in [a, b]$ platí

$$|g(x) - y(x)| < \varepsilon$$

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

Perceptrónová výpočtová jednotka



- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

Perceptróny

Predpokladajme, že trénujeme perceptrón pomocou tréningovej množiny - dvojice

$$\{(\mathbf{x}^1, d^1), (\mathbf{x}^2, d^2), \dots, (\mathbf{x}^m, d^m)\},$$

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

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

Perceptróny

Správna klasifikácia

Ciel - navrhnúť perceptrón, ktorý pre každý vstupný vektor $\mathbf{x}^k$ tréningovej množiny dá výstup y^k , ktorý sa zhoduje s cieľom d^k .

Teda

$$y^k = \text{sgn}(\mathbf{w}^T \mathbf{x}^k) = d^k, k = 1, 2, \dots, m.$$

$$\begin{cases} (\mathbf{x}^k)^T \mathbf{w}^* > 0 & \text{ak } d^k = +1 \\ (\mathbf{x}^k)^T \mathbf{w}^* < 0 & \text{ak } d^k = -1 \end{cases}$$

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

Perceptróny

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

$$\begin{cases} \mathbf{w}^1 - \text{náhodné} \\ \mathbf{w}^{k+1} = \mathbf{w}^k + \rho(d^k - y^k)\mathbf{x}^k, \quad k = 1, 2, \dots \end{cases}$$

kde ρ je kladná konštanta - **učiaci pomer**.

Poznámky:

- ▶ ak riešenie existuje, potom je dosiahnuté po konečnom počte iterácií;
- ▶ vo všeobecnosti táto metóda poskytuje nekonečný počet riešení; perceptrónové učiace pravidlo konverguje k jednému z nich;
- ▶ riešenie je citlivé na hodnotu učiaceho pomeru ρ , ktorá koriguje presentáciu dvojice;
- ▶ počet korekcií k_0 závisí od výberu počiatočných hodnôt vektora váh $\mathbf{w}^1$;
- ▶ ak riešenie neexistuje, perceptrón to neodhalí
- ▶ vektor riešenia existuje, ak daná tréningová množina je lineárne separovateľná

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

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.

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

Najpoužívanéjšie modely NS

Definícia: *Architektúra siete* je konečný vrcholovo- aj hranovo-ohodnotený graf $G = (V, E)$, kde V je konečná množina vrcholov, $|V| = N$, $E, E \subseteq V \times V$ je množina hrán, N nazývame veľkosť siete.

Definícia: *Neurónová sieť* je $\mathcal{N} = (G, Inp, Out, w, c, b, x_{init})$, kde G je architektúra siete, vo vrcholoch sa nachádzajú neuróny, orientované hrany vyjadrujú smer prenosu signálu, $Inp, Out \subseteq V$ sú vstupné, výstupné neuróny, $w : E \rightarrow R$ je ohodnotenie hrán - váhy, $c : V \rightarrow R$ sú prahové hodnoty neurónov, $b : Inp \rightarrow R$ sú váhy vstupov, $x_{init} : V - Inp \rightarrow R$ je počiatočný stav siete.

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

Definícia:

Výpočet siete

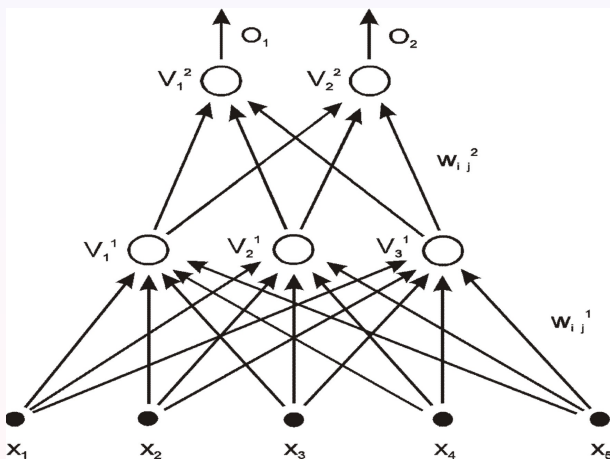
$\mathcal{N}$ v čase $t + 1$ je určený na základe stavov neurónov v čase t

$$x_i(t + 1) = \sigma\left(\sum_{j=1}^N w_{ij}x_j(t) + \sum_{j=1}^M b_ju_j(t) + c_i\right)$$

pre $j = 1, \dots, N$, σ je aktivačná funkcia.

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

Vrstvové siete



Dvojvrstvová sieť

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

Vrstvové siete - Reálne funkcie

Back-propagation algoritmus - najrozšírenejší učiaci algoritmus

Základnou ideou je rozdelenie práce siete:

- ▶ na fázu propagácie vstupného signálu a
- ▶ na fázu propagácie chyby a adaptácie váh.

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

Back-propagation algoritmus

1.fáza- sieť spracúva a prenáša vstupný signál až k výstupnej vrstve, pričom na výpočet aktivačných hodnôt neurónov používa pravidlo

$$V_i^0 = x_i, i = 1, \dots, n$$

$$V_i^m = g(h_i^m) = g \left(\sum_{j=0}^n w_{ij}^m V_j^{m-1} \right)$$

kde $g(x) = \tanh(x)$.

Aktivačné hodnoty neurónov poslednej vrstvy sú zároveň výstupmi siete, teda

$$y_i = V_i^M$$

kde M je počet vrstiev siete (okrem vstupnej vrstvy).

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

Na výpočet chyby siete použijeme kvadratickú chybovú funkciu

$$E = \frac{1}{2} \sum_{i=1}^k [\zeta_i - y_i]^2$$

pričom ζ_i je hodnota očakávaného výstupu i-teho výstupného neurónu a k je počet neurónov vo výstupnej vrstve.

Metóda najväčšieho spádu (gradientu) dáva pravidlo na adaptáciu váh

$$\Delta w_{ij}^m = \frac{\partial E}{\partial w_{ij}^m} = \eta \delta_i^m V_j^{m-1}$$

Parameter η nazývame **učiacim pomerom** (rýchlosť adaptácie), $0 < \eta \leq 1$.

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

2.fázou

práce siete je výpočet hodnôt δ_i^m , potrebných pre adaptáciu. Tie reprezentujú chybu siete, ktorej sa dopustila pri výpočte výstupných hodnôt. Pre výstupnú vrstvu platí

$$\delta_i^M = g'(h_i^M)[\zeta_i - y_i]$$

Postupnou propagáciou chyby naspäť cez sieť získame hodnoty δ_i^m pre všetky vrstvy, pričom

$$\delta_i^{m-1} = g'(h_i^{m-1}) \sum_j w_{ji}^m \delta_j^m$$

Tie potom použijeme pri výpočte nových hodnôt váh prepojení.

- └ Neurón a jeho model, čo dokáže jeden neurón
- └ Polynomiálne prahové jednotky (PTG)

Použitie dopredných neurónových sietí

- ▶ Aproximácie funkcií
- ▶ Predikcia hodnôt
- ▶ Klasifikátory

Funkcionálne priestory, normy

Najdôležitejšie priestory pre nás sú:

1. Priestor všetkých spojitých funkcií nad kompaktným priestorom X (označený $C(X)$) so supremovou normou:

$$\|f\| = \sup_{x \in X} |f(x)|.$$
2. Priestor $\mathcal{L}_p$ všetkých funkcií, pre ktoré je $|f|^p$ integrovateľná na množine X , s tzv. $\mathcal{L}_p$ normou:

$$\|f\| = \left[\int |f(x)|^p dx \right]^{\frac{1}{p}}$$

- .
3. I^1 je interval $\langle 0, 1 \rangle$,

$$I^n = \underbrace{\langle 0, 1 \rangle \times \cdots \times \langle 0, 1 \rangle}_{n\text{-krát}}.$$

Funkcionálne priestory, normy

Pracujeme v priestore R^n a za kompaktnú podmnožinu X budeme považovať jednotkovú kocku, I^n .

Medzi normou a metrikou je indukovaný vzťah $\rho(x, y) = \|x - y\|$.

- ▶ V priestore so supremovou normou je meraná vzdialenosť dvoch funkcií ako supremum rozdielu ich funkčných hodnôt, teda záleží na prípadných lokálnych výkyvoch.
- ▶ U $\mathcal{L}_p$ normy je počítaný integrál ich rozdielu. Je tu brané do úvahy globálne chovanie sa oboch funkcií.

Funkcionálne priestory, normy

- ▶ Existujú funkcie, ktoré sú blízko, keď meriame ich vzdialenosť $\mathcal{L}_p$ normou, ale ďaleko pri supremovej norme.
- ▶ Opačne ale platí, ak sú blízke podľa supremovej normy, tak sú blízko aj podľa $\mathcal{L}_p$ normy.

Výpočtové schopnosti množiny funkcií

- ▶ Vlastnosť **univerzálnej aproximácie** množiny funkcií na vyjadrenie ich výpočtových schopností.
- ▶ Množina funkcií vypočítateľných sieťami s touto vlastnosťou je hustá alebo inými slovami povedané, že jej uzáver je rovný množine všetkých spojitých alebo merateľných funkcií.

Aproximátor, univerzálny aproximátor

Definition

Nech U je trieda funkcií, $T \subseteq U$ a ρ je metrika na U . Triedu T budeme nazývať **aproximátorom** vzhľadom na (U, ρ) , ak T je hustá v U (vzhľadom na topológiu indukovanú metrikou ρ). Ak je trieda funkcií T aproximátorom vzhľadom k triede spojitých reálnych funkcií, budeme ju nazývať **univerzálny aproximátor**.

Prvou prácou, v ktorej je ukázané, že dopredné neurónové siete sú univerzálnymi aproximátormi, je práca R. Hecht - Nielsena z roku 1989.

Jeden rozmer stačí

Theorem

(Kolmogorov, 1957): *Pre ľubovoľné $n \geq 2$ existujú spojité reálne funkcie $\psi^{p,q}$ definované na $I^1 = \langle 0, 1 \rangle$ také, že ľubovoľnú spojitú reálnu funkciu n -premených je možné vyjadriť v tvare*

$$f(x_1, x_2, \dots, x_n) = \sum_{q=1}^{2n+1} \chi^q(\sum_{p=1}^n \psi^{p,q}(x_p)) \quad (4)$$

kde $\chi^q(y)$, $q = 1, 2, \dots, 2n + 1$ sú spojité funkcie jednej premennej.

Dôkaz je konštruktívny, veľa hovorí o vlastnostiach funkcií $\psi^{p,q}$ a χ^q .

Základné myšlienky dôkazu

1. Skoro celý I^n pokryjeme navzájom disjunktnými n -rozmernými kockami.
2. Potom zostrojíme spojitú, monotónne rastúcu funkciu ψ^{pq} , ktoré sú definované na I^1 tak, aby pre dané q , funkcia n -premenných $\sum_{p=1}^n \psi^{pq}(x_p)$ zobrazila n -rozmerné kocky do navzájom disjunktných intervalov.
3. Na týchto intervaloch definujeme funkcie χ_r^q podľa predpisu a mimo intervalov dodefinujeme na spojitú funkcie.
4. Ukážeme, že funkcie χ^q , ktoré potrebujeme, sú funkcie $\chi^q = \lim_{r \rightarrow \infty} \chi_r^q$.

Theorem

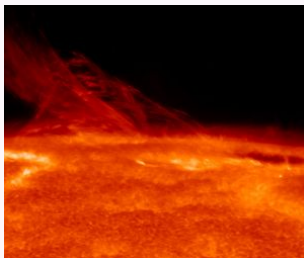
(Kurková, 1991) *Nech $n \geq 2$ je prirodzené číslo, $\sigma : R \rightarrow [0, 1]$ je schodovitá funkcia, $f : [0, 1] \rightarrow R$ je spojitá funkcia a ϵ je kladné reálne číslo. Potom existuje prirodzené číslo k a schodovité funkcie typu $\sigma \psi^{p,i}, \phi^i, i = 1 \dots, k, p = 1, \dots, n$ také, že pre každé $(x_1, \dots, x_n) \in [0, 1]^n$ platí*

$$|f(x_1, \dots, x_n) - \sum_{i=1}^k \phi^i \left(\sum_{p=1}^n \psi^{p,i}(x_p) \right)| < \epsilon. \quad (5)$$

Nový cyklus Slnka prinesie búrky

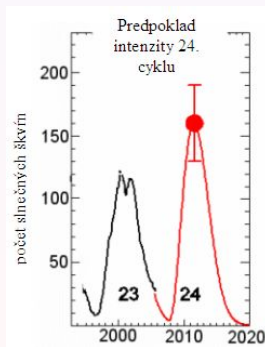
DAGMAR GREGOROVÁ PRE SME.SK, 9. 1. 2008

Po niekoľkých mesiacoch minimálnej slnečnej aktivity sa v štvrtý deň nového roka objavil na žiarivom povrchu našej hviezdy malý fliačik. Nič dramatické. A predsa táto slnečná škvrna vzbudila pozornosť astronómov.



Začiatok nového slnečného cyklu

Astronómom, ktorí sa pozreli na obraz magnetických polí bolo hneď jasné. Vraví sa o jedenásťročnom cykle, ale to je len jeho priemerná hodnota a v skutočnosti Slnko nedodržiava presný grafikon, občas si ho aj o vyše roka skrúti, či predĺži.



Čo to je geomagnetická búrka (GMB)

- ▶ magnetické pole Zeme - pole veľmi veľkého trvalého magnetu,
- ▶ kvantitatívny popis
 - ▶ fyzikálna veličina - **magnetická indukcia B** , okolo $0,5 \cdot 10^{-4} \text{T}$
 - ▶ fyzikálna veličina - **intenzita magnetického poľa H** , približne $24 \text{A} \cdot \text{m}^{-1}$
- ▶ sila a smer magnetického poľa Zeme sa mení
- ▶ dochádza k výkyvom, ktoré trvajú niekoľko hodín alebo dní - **geomagnetické búrky (GMB)**

Prečo vznikajú GMB

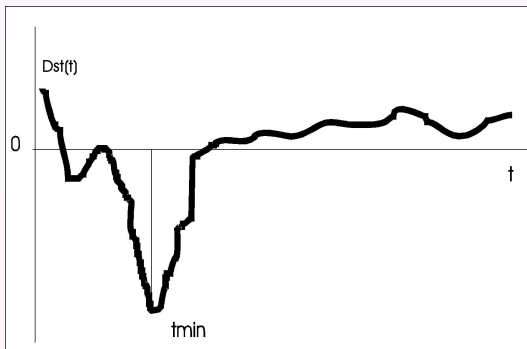
- ▶ procesy prebiehajúce na Slnku
- ▶ slnečné erupcie spôsobujú výrony častíc a žiarenia do medziplanetárneho priestoru
- ▶ prúd týchto častíc slnečného pôvodu, ktoré letia medziplanetárnym priestorom veľkou rýchlosťou (zo Slnka unikajú rýchlosťou 1000 až 3000 km/s) - **slnečným vetrom**

Čo môžu spôsobiť GMB

- ▶ aj keď pojem geomagnetická búrka znie dramaticky, pre človeka nie je nebezpečná
- ▶ ohrozené sú veľmi dlhé vodivé potrubia a elektrické vedenia najmä v smere poludníkov; meniace sa pole v nich naindukuje napätie, ktoré by mohlo poškodiť zariadenia rozvodných sietí, napríklad transformátory
- ▶ oblaky vysokoenergetických nabitých častíc však môžu spôsobiť vážne problémy tam, kde priestor už nie je pod ochranou štítu geomagnetického poľa - na družiciach, vesmírnych sondách, staniciach aj astronautom

Priebeh indexu D_{st} geomagnetickej búrky

Zmeny magnetického poľa Zeme sú sledované pomocou veličiny D_{st} - **index geomagnetického poľa Zeme** - v kľude ± 20 nT



Funkcionálne závislosti pri skúmaní GMB

Index D_{st} fyzici sledujú v závislosti od veličín:

- ▶ B_z - z-ová zložka vektora intenzity medziplanetárneho magnetického poľa, ktorá nadobúda hodnoty v intervale $\pm 50 nT$,
- ▶ σ_{B_z} - stredná kvadratická odchýlka B_z , ktorá charakterizuje výkyvy tejto veličiny,
- ▶ n - počet častíc slnečného vetra v 1 cm^3 ,
- ▶ v - rýchlosť častíc slnečného vetra, ktorá dosahuje až 1200 km/s .

Dôležitý problém - predikcia GMB

- ▶ index D_{st} - funkciou 4 premenných
- ▶ sledovateľný priebeh v čase
- ▶ z nameraných hodnôt vytvárame aproximujúce funkcie
- ▶ využitie týchto poznatkov na predikciu GMB niekoľko hodín dopredu
- ▶ **predikciou** rozumieme predpovedanie ďalšieho priebehu časovej postupnosti na základe predchádzajúcich hodnôt

Časové postupnosti

Časovou postupnosťou rozumieme množinu hodnôt danej veličiny závislej na čase, ktoré reprezentujú jej priebeh v určitom časovom intervale.

Podľa typu hodnôt časovej postupnosti ich delíme na

- ▶ **statické časové postupnosti** -

$$P = \{f(t_0), f(t_1), \dots, f(t_n); t_0 < t_1 < \dots < t_n\},$$

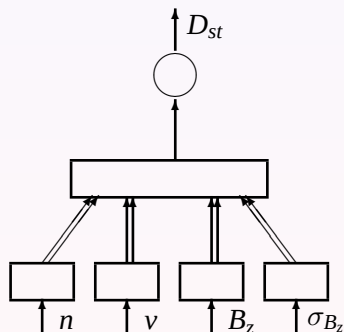
kde f je vyjadrenie závislosti danej veličiny na čase

- ▶ **dynamické časové postupnosti** - daná veličina je spojitou funkciou času

$$P = f(t) \quad \text{pričom } t \in \langle t_0, t_1 \rangle \quad t_0 < t_1$$

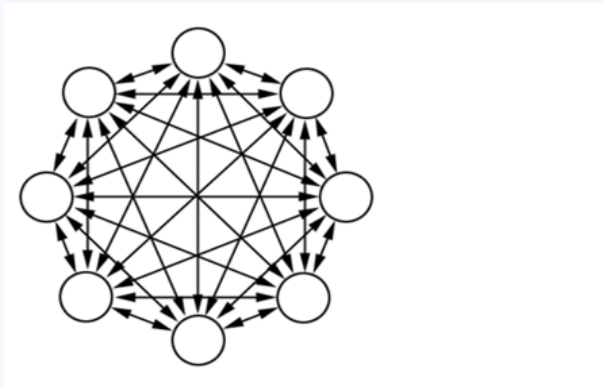
- └ Časové postupnosti
- └ Predikcia GMB pomocou neuronových sietí

Modely neuronových sietí



- └ Časové postupnosti
- └ Predikcia GMB pomocou neurónových sietí

Hopfieldove neurónové siete



Hopfieldová neurónová sieť

Topológia siete

Hopfieldove siete - použitie

- ▶ Asociatívne pamäte
- ▶ Riešenie optimalizačných úloh

Energetická funkcia siete

$$E_{hop} = -\frac{1}{2} \sum_i^n \sum_j^n w_{ij} V_i V_j + \sum_i^n \Theta_i V_i \quad (6)$$

Polynómy viacerých premenných

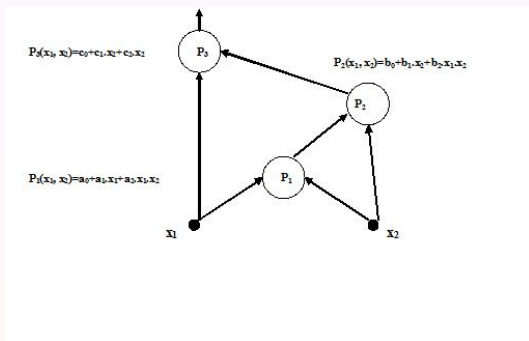
$$P(x_1, x_2, \dots, x_d) = a_0 + \sum_{i=1}^d a_i x_i + \sum_{i=1}^d \sum_{j=1}^d a_{ij} x_i x_j + \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d a_{ijk} x_i x_j x_k + \dots, \quad (7)$$

kde $a_0, a_i, a_{ij}, \dots$ sú koeficienty polynómu, x_i sú vstupné premenné.

Tvar polynómu je v lineárnom tvare - dá sa skonštruovať jednovrstvová neurónová sieť.

Vlastnosti polynómov

- ▶ súčet a rozdiel polynómov je polynóm
- ▶ súčin polynómov je polynóm
- ▶ toto umožňuje použitie neurónov s polynomiálnymi prenosovými funkciami



Trénovanie polynomiálnych neurónových sietí

Poznámky:

- ▶ Pri aproximácii funkcií pomocou polynómov sa používajú rôzne metódy na výpočet koeficientov polynómov.
- ▶ Neurónové siete sú vytvárané práve kvôli ich vlastnosti adaptovať sa na trénovacie údaje.
- ▶ Je zrejmé, že aj pre tieto siete sú použiteľné **back - propagation metódy**, samozrejme prispôbené použitým polynómom v polynómialnej neurónovej sieti.
- ▶ **Otázka: Ktorá sieť je najlepšia?**

Ďakujem za pozornosť