

# Predikcia časových radov pomocu hierarchických neurónových sietí s echo stavmi

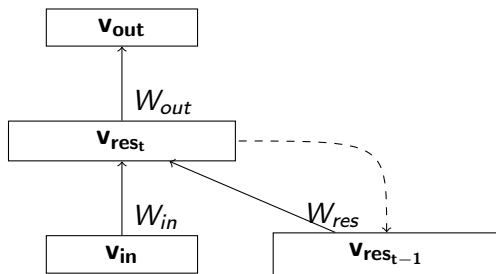
Bc. Viliam Dillinger

pod vedením doc. Ing. Igora Farkaša PhD.

16. marca 2011

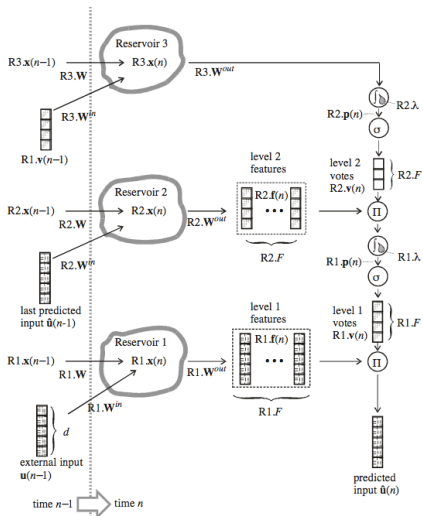
- 1 Neurónové siete s echo stavmi
- 2 Hierarchické neurónové siete s echo stavmi
- 3 Mackey-Glass Time series
- 4 Reber Grammar
- 5 Pokračovanie mojej práce

# Neurónové siete s echo stavmi



- $\mathbf{v}_{res_t}$  je vysokorozmerný rezervoár
- $\mathbf{W}_{res}$  je riedka matica so spektrálnym rádiusom  $< 1$
- $\mathbf{v}_{res_t} = \sigma(\mathbf{v}_{in} \mathbf{W}_{in} + \mathbf{v}_{res_{t-1}} \mathbf{W}_{res})$
- $\mathbf{v}_{out} = \mathbf{v}_{res_t} \mathbf{W}_{out}$
- pri učení sa mení iba  $\mathbf{W}_{out}$

## Hierarchické neuronové siete s echo stavmi



## Inicializácia a parametre

- $Rk.x$  - vysokorozmerný rezervoár  $k$ -tej vrstvy
- $Rk.W$  - riedka matica, spektrálny rádius  $< 1$
- $Rk.f_i$  - riedka matica  $i$ -teho príznaku  $k$ -tej vrstvy (každý príznak vidí len časť rezervoára)
- $Rk.\lambda$  - vote leaky rate  $k$ -tej vrstvy
- $Rk.a$  - reservoir leaky rate  $k$ -tej vrstvy
- $Rk.\gamma$  - rýchlosť učenia sa  $k$ -tej vrstvy
- Všetky matice sa inicializujú na náhodné hodnoty medzi  $(-1, 1)$
- $Rk.x$  a  $Rk.v$  sa inicializujú na hodnoty medzi  $(-0.01, 0.01)$
- $Rk.i$  - vstup  $k$ -tej vrstvy

$$R1.i(n) = input(n)$$

$$R(k+1).i(n) = Rk.v(n-1)$$

# Šírenie signálu

- Zmena stavu rezervoáru  $k$ -tej úrovne:

$$Rk.x(n) = (1 - Rk.a)Rk.x(n-1) + \sigma (Rk.WRk.x(n-1) + Rk.W^{in}Rk.i(n-1))$$

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

- Abstrakcia  $i$ -teho príznaku:

$$Rk.f_i(n) = Rk.W_i^{out} [Rk.x(n); Rk.i(n)]$$

- Výstup z vrstvy:

$$Rk.v(n) = \sigma (\mathcal{L}_{Rk.\lambda} (R[k+1].f(n)R[k+1].v(n)))$$

$$\mathcal{L}_\lambda(q(n)) = (1 - \lambda)\mathcal{L}_\lambda(q(n-1)) + \lambda q(n)$$

- výstup najvrchnejšej vrstvy, má vektor hlasov  $\mathbf{v}(n) = (1)$
- výstup najnižšej vrstvy neprechádza cez sigmoid ani leaky integration a je výstupom celého modelu.

# Učenie sa

- Supervised learning (učenie sa s učiteľom)
- menia sa len matice príznakov -  $Rk.f_j$
- Chyba:

$$E_1(n) = target(n) - R1.v(n)$$

$$E_k(n) = R[k-1].f^T(n)E_{k-1}(n) * R[k-1].\lambda\sigma'(R[k-1].p(n))$$

- Učiace sa pravidlo:

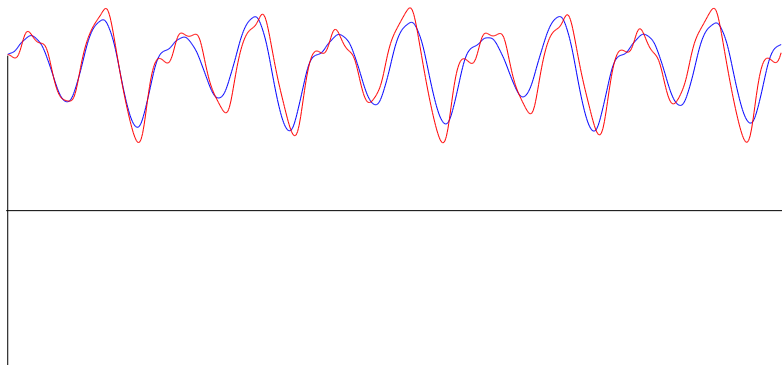
$$Rk.W_i^{out}(n+1) = Rk.W_i^{out}(n) + Rk.\gamma Rk.v_{[j]}(n)E_k(n) [Rk.x(n); Rk.i(n)]^T$$

# Dáta

- $\frac{ds(t)}{dt} = \alpha \frac{s(t-\tau)}{1+s^{10}(t-\tau)} - \beta s(t)$
- premenné (v matlabe mgdata.dat):
  - $\alpha = 0.2$
  - $\beta = 0.1$
  - $s(0) = 1.2$  (pre  $t < 0$   $s(t) = 0$ )
  - $\tau = 17$
- vstup je vektor  $[s(t-18), s(t-12), s(t-6), s(t)]$
- predikcia  $s(t+85)$
- trénovacia množina 500 vektorov
- testovacia množina 500 vektorov (disjunktné)



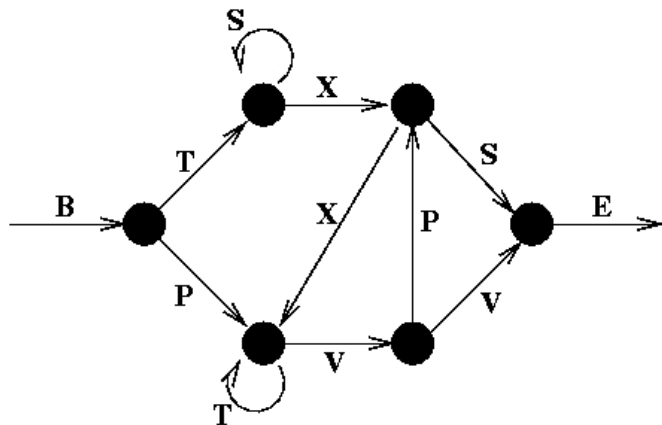
## Výsledky



# Výsledky

- $$NRMSE = \sqrt{\frac{\sum_{k=1}^n (y_k - \hat{y}_k)^2}{\sum_{k=1}^n (y_k - \bar{y}_k)^2}}$$
- $$\bar{y} = \frac{\sum_{k=1}^n y_k}{k}$$
- najlepšie mnou dosiahnutý výsledok –  $NRMSE = 0.378$
- J. González et al. (2006) dosiahli  $NRMSE = 0.0855$ , najhorši výsledok s ktorým sa porovnávali  $NRMSE = 0.378$  (Resource Allocation Network)

## Reber Grammar



# Trénovanie a výsledky

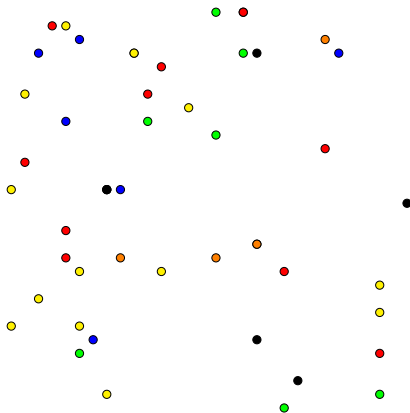
- v každom stave rovnaká pravdepodobnosť vygenerovania oboch písmen
- pri trénovaní sieť dostala na vstup aktuálny znak a mala predikovať ďalší symbol
- one-hot kódovanie
- po natrénovaní s približne 50% pravdepodobnosťou vedela určiť ďalšie písmenko
- po normalizácii výstupného vektoru aproximovala prechodovú funkciu (vcelku úspešne)

# Aproximácia stavu a prechodovej funkcie

- 2 vrstvy - na prvej niekoľko malých sietí (aby nedokázali túto úlohu riešiť samostatne), na druhej jedna silná
- hypotéza - výstup vrchnej vrstvy - aproximácia stavu, výstup spodnej vrstvy - aproximácia prechodovej funkcie
- vizualizácia pomocou SOM
- pri jednoduchých automatoch pravdivá, pri zložitejších nie.

Jednoduchý automat ( $abc^n$ )

# Reber Grammar



# Pokračovanie mojej práce

- Mackey-Glass Time series
  - pokúsiť sa prekonať chybu, ktorú dosiahli odborníci z univerzity v Granade (alebo sa aspoň k tomu priblížiť)
- Reber Grammar a iné symbolové dynamiky
  - pomocou SOM sa pokúsiť abstrahovať stav predikovaného automatu zo siete
  - ukázať, že sieť vie generalizovať (trénovacia a testovacia množina)
  - zopakovať niektoré pokusy ako Cleeremans, Servan-Schreiber a McClelland (1989)
- Otestovať správanie sa siete, ak na jednej úrovni nebudú rovnako silné siete



Ďakujem za pozornosť. Nech sa páči, priestor pre Vaše otázky.