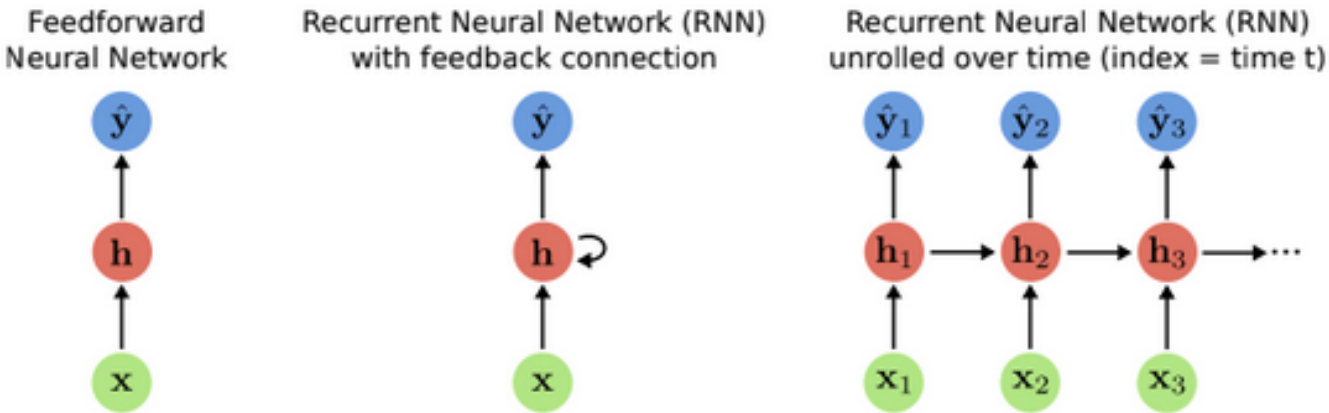
Co jsme probírali minule

Konvoluční neuronová síť

1. Operace konvoluce
2. Konvoluční vrstva
3. Architektura konvoluční neuronové sítě
4. Známé modely konvolučních neuronových sítí (úvod)
5. Známé modely konvolučních neuronových sítí (dokončení)
6. Učení konvolučních neuronových sítí: přenesené učení, regularizace, předzpracování dat
7. Varianty a aplikace konvolučních neuronových sítí (přehled)

Rekurentní neuronová síť (RNN)

Jednoduchá rekurentní síť (vanilla RNN)

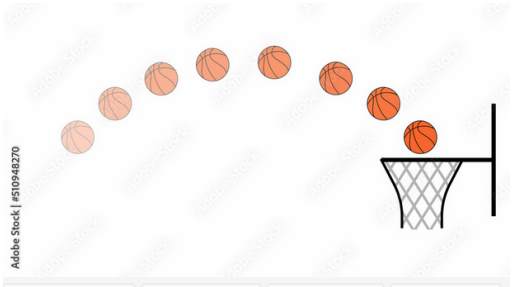


- zobecnění vrstevnaté neuronové sítě o vnitřní paměť
- myšlenka: neurony si udržují svůj vnitřní stav
- výstup neuronu závisí nejen na vstupních datech, ale i na jeho stavu (minulém výstupu)
- využití: zpracování posloupností, časových řad ap.
- neuronům v RNN se někdy říká „buňky" (*memory cells*)

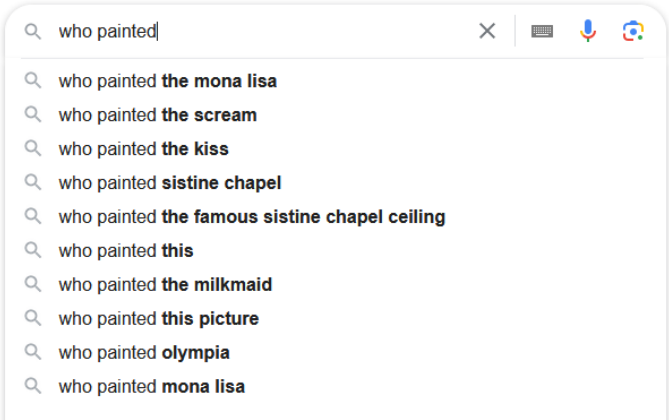
Rekurentní neuronová síť (RNN)

Příklady úloh:

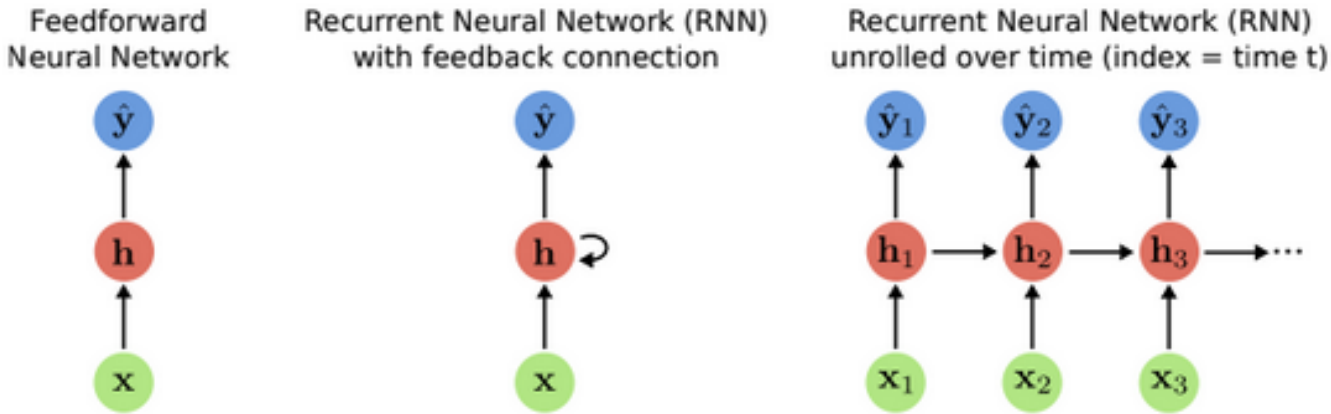
- jaká bude poloha hozeného míče v dalším okamžiku?



- dokonči větu:

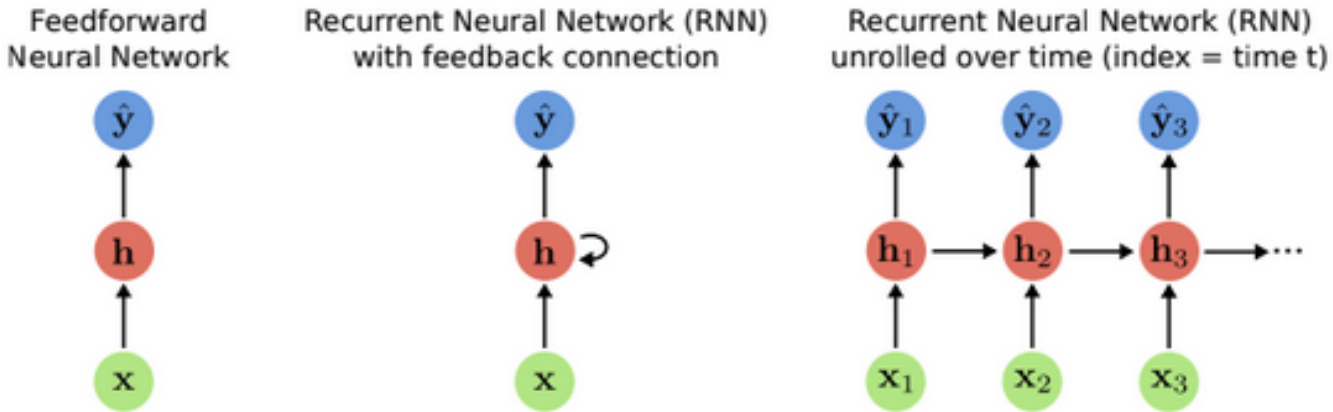


Jednoduchá rekurentní síť (vanilla/Elman RNN)



- myšlenka: neuron si udržuje svůj vnitřní stav: $h_t = f_h(h_{t-1}, x_t)$
- výstup neuronu: $y_t = f_y(h_t)$
- modelu předložíme celou posloupnost dat
  - libovolné délky
  - model si můžeme ,,rozbalit" (obrázek vpravo)
  - f_y, f_h se (pro jednu vstupní posloupnost) v čase nemění
- výhoda: je možné zpracovat posloupnosti libovolné délky
- pro novou posloupnost se stav resetuje

Jednoduchá rekurentní síť (vanilla/Elman RNN)

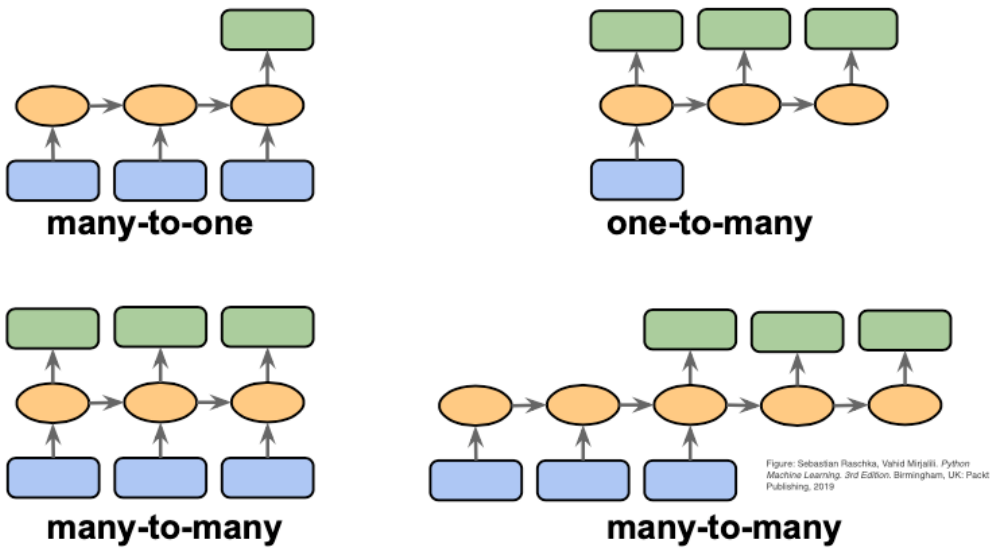


- vnitřní stav: $h_t = tanh(w_h h_{t-1} + w_x x_t + b_h)$
- výstup neuronu: $y_t = w_y h_t + b_y$
- obvyklá přenosová funkce je tanh
- modelu předložíme celou posloupnost dat
  - model si můžeme ,,rozbalit" (obrázek vpravo)
  - váhy a biasy w, b se (pro jednu vstupní posloupnost) v čase nemění

Rekurentní neuronová síť (RNN)

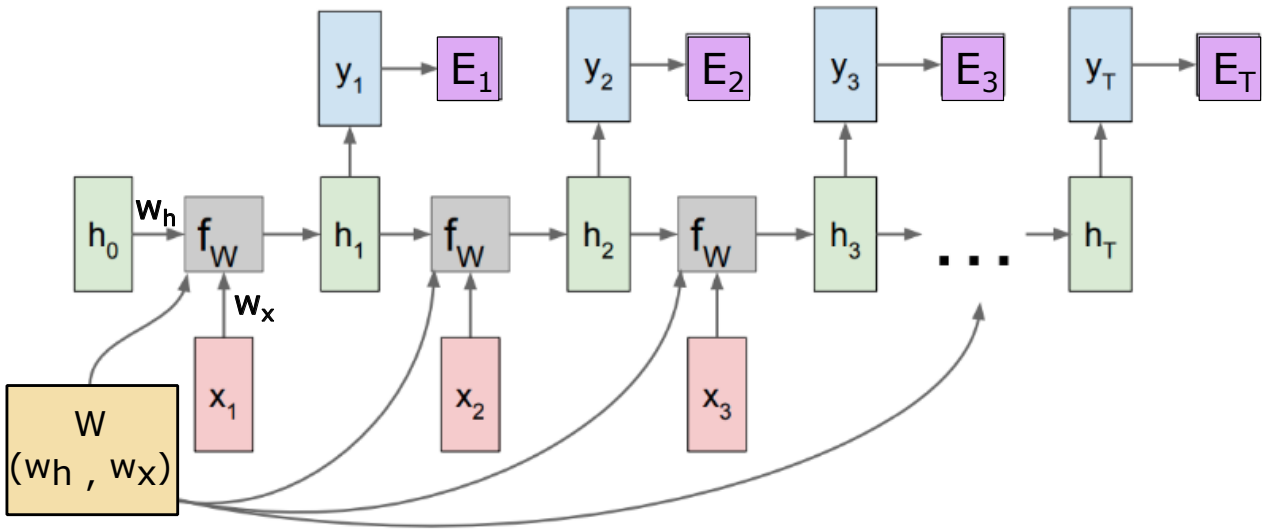
Typy úloh nad sekvenčními daty

- one-to-one - např. obyčejná klasifikace
- many-to-one - např. klasifikace sentimentu, rozpoznání akce na videu
- one-to-many - např. vygenerovat slovní popis k obrázku, sentence tagging
- many-to-many - strojový překlad, object tracking



Rekurentní neuronová síť (RNN)

Algoritmus učení: backpropagation through time (BPTT)

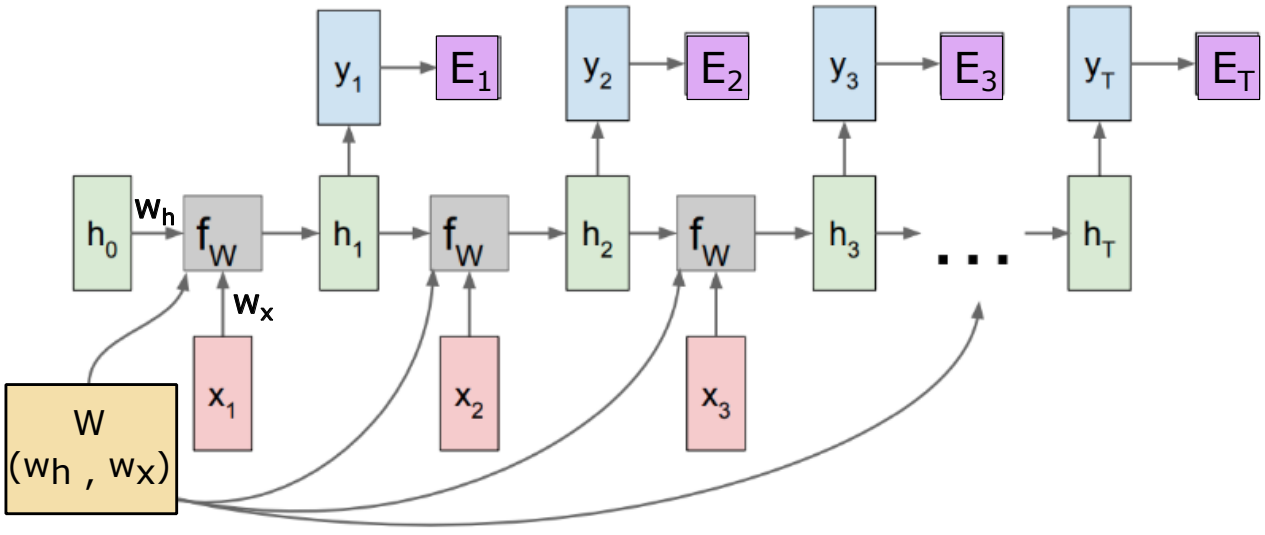


- ukázka pro architekturu many-to-many
- zpětná propagace pro celou vstupní posloupnost ,,najednou"

$$h_t = f_W(h_{t-1}, x_t) = tanh(w_h h_{t-1} + w_x x_t + b_h)$$

Rekurentní neuronová síť (RNN)

Algoritmus učení: backpropagation through time (BPTT)



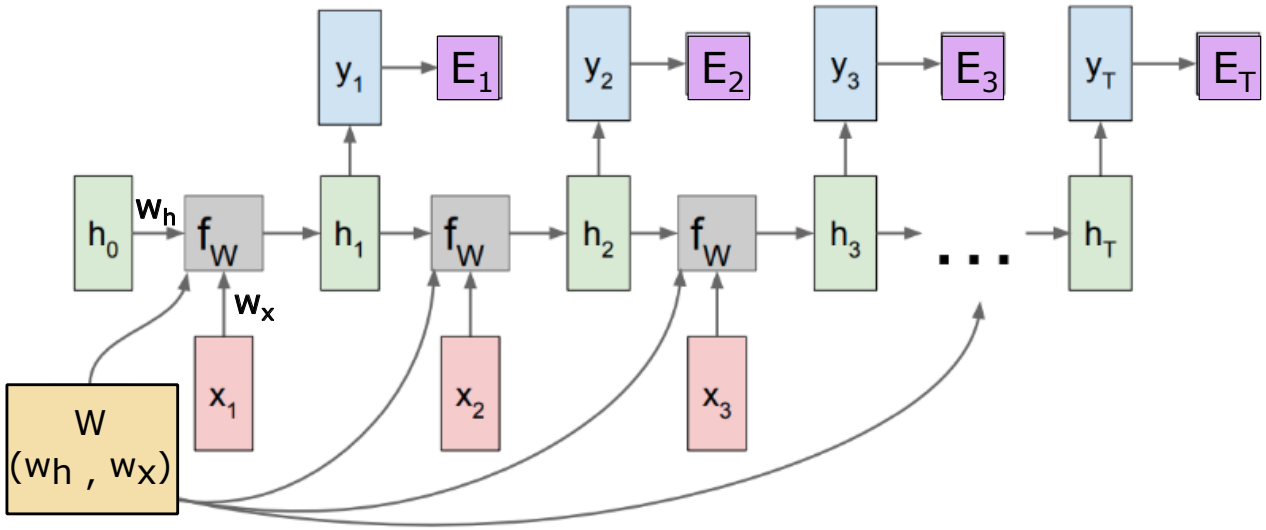
- chyba se propaguje v čase:

$$\frac{\partial E_1}{\partial w_x} = \frac{\partial E_1}{\partial h_1} \cdot \frac{\partial h_1}{\partial w_x}$$

$$\frac{\partial E_2}{\partial w_x} = \frac{\partial E_2}{\partial h_2} \cdot \frac{\partial h_2}{\partial w_x} + \frac{\partial E_2}{\partial h_2} \cdot \frac{\partial h_2}{\partial h_1} \cdot \frac{\partial h_1}{\partial w_x}$$

Rekurentní neuronová síť (RNN)

Algoritmus učení: backpropagation through time (BPTT)

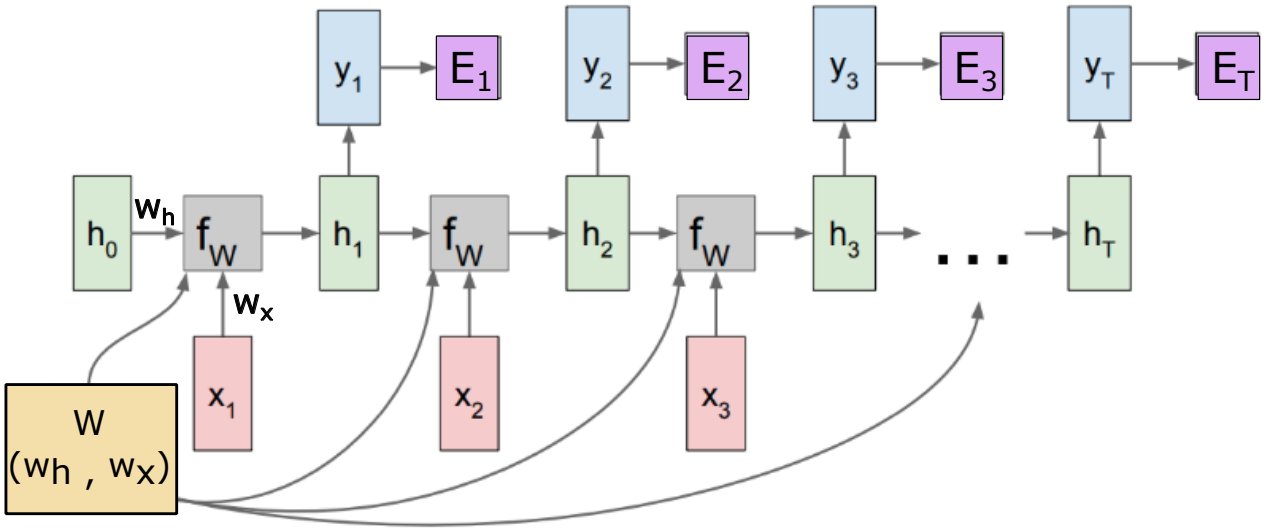


- chyba se propaguje v čase:

$$\frac{\partial E_3}{\partial w_x} = \frac{\partial E_3}{\partial h_3} \cdot \frac{\partial h_3}{\partial w_x} + \frac{\partial E_3}{\partial h_3} \cdot \frac{\partial h_3}{\partial h_2} \cdot \frac{\partial h_2}{\partial w_x} + \frac{\partial E_3}{\partial h_3} \cdot \frac{\partial h_3}{\partial h_2} \cdot \frac{\partial h_2}{\partial h_1} \cdot \frac{\partial h_1}{\partial w_x}$$

Rekurentní neuronová síť (RNN)

Algoritmus učení: backpropagation through time (BPTT)



- celkově:

$$\frac{\partial E}{\partial w_x} = \frac{1}{n}\frac{\partial E_n}{\partial w_x} = \frac{1}{n}\sum_{t=1}^{n}\sum_{k=1}^{t}\frac{\partial E_t}{\partial h_t}\left(\prod_{j=k+1}^{t}\frac{\partial h_j}{\partial h_{j-1}}\right)\frac{\partial h_k}{\partial w_x}$$

Rekurentní neuronová síť (RNN)

Problémy algoritmu BPTT

- problém mizejících gradientů:

$$\frac{\partial E}{\partial w_x} = \frac{1}{n}\frac{\partial E_n}{\partial w_x} = \frac{1}{n}\sum_{t=1}^{n}\sum_{k=1}^{t}\frac{\partial E_t}{\partial h_t}\left(\prod_{j=k+1}^{t}\frac{\partial h_j}{\partial h_{j-1}}\right)\frac{\partial h_k}{\partial w_x}$$

$$h_t = tanh(w_h h_{t-1} + w_x x_t + b_h)$$

  - pokud jsou derivace přenosové funkce blízké 0 $\rightarrow$ násobíme hodně malých čísel $\rightarrow$ moc malý krok učení
- problém explodujících gradientů
  - pokud jsou derivace přenosové funkce moc velké $\rightarrow$ násobíme hodně velkých čísel $\rightarrow$ moc velký krok učení

$\rightarrow$ model má problém se učit

$\rightarrow$ model zapomíná, má špatnou dlouhodobou paměť

Řešení problému mizejících explodujících gradientů

Gradient clipping

- řeší problém explodujících gradientů
- pokud je norma gradientu větší než daná mez, normalizujeme ho (vydělíme normou)

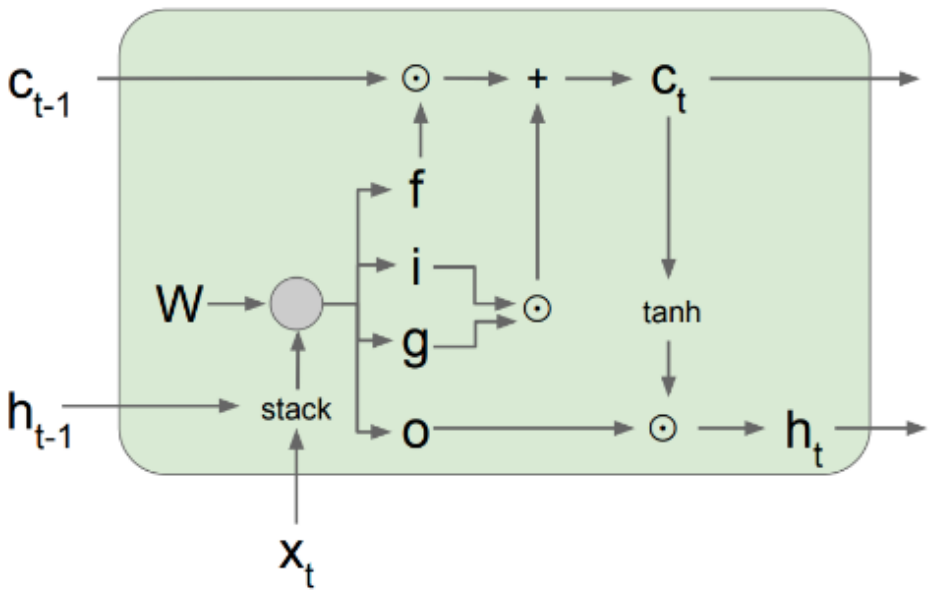
Problém mizejících gradientů

- model zapomíná, má špatnou dlouhodobou paměť
- ukázka obtížného příkladu: ,,The dogs in the neighborhood are ..." (barking)
- **řešení**:
  - echo state networks
  - LSTM (Long Short Term Memory)
  - GRU (Gated recurrent unit)

LSTM (Long Short Term Memory)

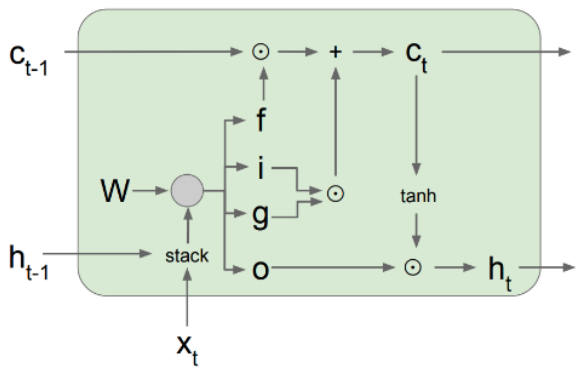
Myšlenka:

- násobení nahradíme sčítáním
- krátkodobý stav buňky: h_t
- dlouhodobý stav buňky: c_t uchovává dlouhodobou informaci
- používá brány (gates): vstupní (i), zapomínací (f), výstupní (o), brána aktuálního stavu (g)



LSTM (Long Short Term Memory)

- **vstupní brána: zda zapsat** $i_t = \sigma(w_h^{(i)} h_{t-1} + w_x^{(i)} x_t + b_h^{(i)})$
- **brána aktuálního stavu: kolik zapsat**
  $g_t = tanh(w_h^{(g)} h_{t-1} + w_x^{(g)} x_t + b_h^{(g)})$
- **zapomínací brána: zda zapomenout**
  $f_t = \sigma(w_h^{(f)} h_{t-1} + w_x^{(f)} x_t + b_h^{(f)})$
- **výstupní brána: kolik dát na výstup**
  $o_t = \sigma(w_h^{(o)} h_{t-1} + w_x^{(o)} x_t + b_h^{(o)})$



Rekurentní neuronová síť (RNN) - shrnutí

- RNN umožňují zpracovávat vstupní vzory s různou délkou
- variabilní architektura (one-to-many, many-to-one,...)
- vhodné pokud jsou v datech vnitřní závislosti dané posloupností (např. časem)
- jednoduché RNN se obtížně učí $\rightarrow$ v praxi se používají LSTM a GRU, gradient clipping
- architektura je ,,hluboká", ale v jiném - ,,časovém" rozměru (ale existují i vícevrstvé rekurentní sítě)
- existuje celá řada dalších architektur RNN