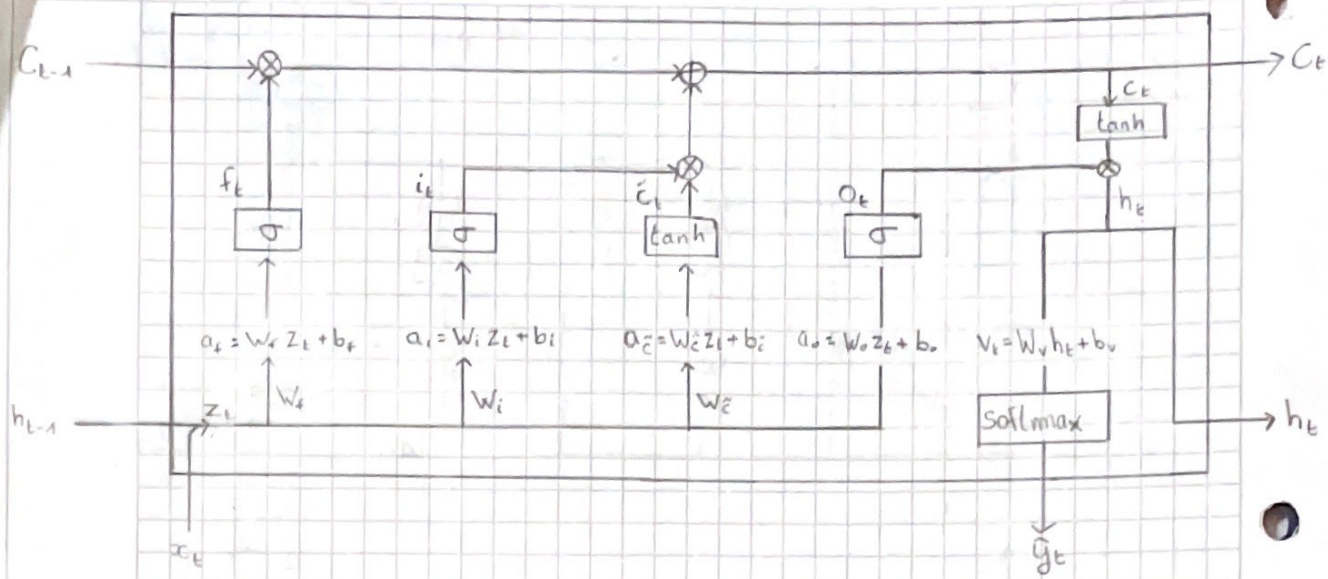
Lstm.



$$\sigma(x) = \frac{1}{1+e^{-x}} \quad , \quad \text{sigmoid}(\vec{x})_i = \frac{e^{x_i}}{\sum_{j=1}^{|\vec{x}|} e^{x_j}}$$

$\otimes$: $A, B \in \mathbb{C}^{m\times n}$; $A \otimes B \in \mathbb{C}^{m\times n}$, $(A\otimes B)_{i,j} = (A)_{i,j} \times (B)_{i,j}$

$\oplus$: $A, B \in \mathbb{C}^{m\times n}$; $A \oplus B \in \mathbb{C}^{m\times n}$, $(A\oplus B)_{i,j} = (A)_{i,j} + (B)_{i,j}$

: Concatenation de vecteur   : copy

<u>Propagation avant :</u>

$z = [h_{t-1}, x_t]$

$a_f = W_f z_t + b_f$

$f_t = \sigma(a_f)$

$a_i = W_i z_t + b_i$

$i_t = \sigma(a_i)$

$a_{\tilde{c}} = W_{\tilde{c}} z_t + b_{\tilde{c}}$

$\tilde{C}_t = \tanh(a_{\tilde{c}})$

$a_o = W_o z_t + b_o$

$O_t = \sigma(a_o)$

$C_t = C_{t-1} \otimes f_t + i_t \otimes \tilde{C}_t$

$h_t = \tanh(C_t) \otimes O_t$

$v_t = W_v h_t + b_v$

$\hat{y}_t = \text{softmax}(v_t)$

Retropropagation du gradient :

Soit L_k la loss au temps k : $L_k = - \sum_{t=k}^{T} \sum_{j=1}^{|V|} y_{t,j} \log(\hat{y}_{t,j})$

Soit $L = L_1$ la loss totale

$$\frac{\partial L}{\partial v_{t,i}} = \frac{\partial}{\partial v_{t,i}} \left(- \sum_{t'=1}^{T} \sum_{j=1}^{|V|} y_{t',j} \log(\hat{y}_{t',j}) \right)$$

$$= \frac{\partial}{\partial v_{t,i}} \left(\sum_{j=1}^{|V|} y_{t,j} \log(\hat{y}_{t,j}) \right)$$

$$= \frac{\partial}{\partial v_{t,i}} \left(- \sum_{j=1}^{|V|} y_{t,j} \left(v_{t,j} - \log\left(\sum_{k=1}^{|V|} e^{v_{t,k}} \right) \right) \right)$$

$$= - \sum_{j=1}^{|V|} y_{t,j} \frac{\partial v_{t,j}}{\partial v_{t,i}} + \sum_{j=1}^{|V|} y_{t,j} \frac{\partial}{\partial v_{t,i}} \log\left(\sum_{k=1}^{|V|} e^{v_{t,k}} \right)$$

$$= - y_{t,i} + \sum_{j=1}^{|V|} y_{t,j} \frac{e^{v_{t,i}}}{\sum_{k=1}^{|V|} e^{v_{t,k}}}$$

$$= y_{t,i} + \hat{y}_{t,i} \left(\sum_{j=1}^{|V|} y_{t,j} \right)$$

La sortie est un vecteur one-hot encoded, ainsi $\sum_{j=1}^{|V|} y_{t,j} = 1$

$$= - y_{t,i} + \hat{y}_{t,i} \Rightarrow \boxed{\frac{\partial L}{\partial v_t} = \hat{y}_t - y_t}$$

$$\frac{\partial L}{\partial h_t} = \frac{\partial L}{\partial v_t} \frac{\partial v_t}{\partial h_t} + \frac{\partial L}{\partial h_{t+1}}$$

$$= \frac{\partial L}{\partial v_t} \frac{\partial}{h_t} \left(W_v h_t + b_v \right) + \frac{\partial L}{\partial h_{t+1}}$$

$$\Rightarrow \boxed{\frac{\partial L}{\partial h_t} = W_v^T \frac{\partial L}{\partial v_t} + \frac{\partial L}{\partial h_{t+1}}}$$

$$\frac{\partial L}{C_t} = \frac{\partial L}{\partial h_t} \frac{\partial h_t}{\partial C_t} + \frac{\partial L}{\partial C_{t+1}}$$

$$= \frac{\partial L}{\partial h_t} \frac{\partial}{\partial C_t} \left(\tanh(C_t) \otimes O_t \right) + \frac{\partial L}{\partial C_{t+1}}$$

$$\Rightarrow \boxed{\frac{\partial L}{\partial C_t} = \frac{\partial L}{\partial h_t} \left((1 - \tanh^2(C_t)) \otimes O_t \right) + \frac{\partial L}{\partial C_{t+1}}}$$

$$\frac{\partial L}{\partial o_t} = \frac{\partial L}{\partial h_t} \frac{\partial h_t}{\partial o_t}$$

$$= \frac{\partial L}{\partial h_t} \frac{\partial}{\partial o_t} \left(\tanh(C_t) \otimes o_t \right)$$

$$\Rightarrow \boxed{\frac{\partial L}{\partial o_t} = \frac{\partial L}{\partial h_t} \tanh(C_t)}$$

$$\frac{\partial L}{\partial a_o} = \frac{\partial L}{\partial o_t} \frac{\partial o_t}{\partial a_o}$$

$$= \frac{\partial L}{\partial o_t} \frac{\partial}{\partial a_o} \left(\sigma(a_o) \right)$$

$$\Rightarrow \boxed{\frac{\partial L}{\partial a_o} = \frac{\partial L}{\partial o_t} \left(\sigma(a_o)(1 - \sigma(a_o)) \right)}$$

$$\frac{\partial L}{\partial \tilde{C}_t} = \frac{\partial L}{\partial C_t} \frac{\partial C_t}{\partial \tilde{C}_t}$$

$$= \frac{\partial L}{\partial C_t} \frac{\partial}{\partial \tilde{C}_t} \left(C_{t-1} \otimes f_t + i_t \otimes \tilde{C}_t \right)$$

$$\Rightarrow \boxed{\frac{\partial L}{\tilde{C}_t} = \frac{\partial L}{\partial C_t} \left(C_{t-1} \otimes f_t + i_t \right)}$$

$$\frac{\partial L}{\partial a_{\tilde{c}}} = \frac{\partial L}{\partial \tilde{C}_t} \frac{\partial \tilde{C}_t}{\partial a_{\tilde{c}}}$$

$$= \frac{\partial L}{\partial \tilde{C}_t} \frac{\partial}{\partial a_{\tilde{c}}} \left(\tanh(a_{\tilde{c}}) \right)$$

$$\Rightarrow \boxed{\frac{\partial L}{\partial a_{\tilde{c}}} = \frac{\partial L}{\partial \tilde{C}_t} \left(1 - \tanh^2(a_{\tilde{c}}) \right)}$$

$$\frac{\partial L}{\partial i_t} = \frac{\partial L}{\partial C_t} \frac{\partial C_t}{\partial i_t}$$

$$= \frac{\partial L}{\partial C_t} \frac{\partial}{\partial i_t} \left(C_{t-1} \otimes f_t + i_t \otimes \tilde{C}_t \right)$$

$$\boxed{\frac{\partial L}{\partial i_t} = \frac{\partial L}{\partial C_t} \left(C_{t-1} \otimes f_t + \tilde{C}_t \right)}$$

$$\frac{\partial L}{\partial a_i} = \frac{\partial L}{\partial i_t} \frac{\partial i_t}{\partial a_i}$$

$$= \frac{\partial L}{\partial i_t} \frac{\partial}{\partial a_i} \left(\sigma(a_i) \right)$$

$$\Rightarrow \boxed{\frac{\partial L}{\partial a_i} = \frac{\partial L}{\partial i_t} \left(\sigma(a_i) \left(1 - \sigma(a_i) \right) \right)}$$

$$\frac{\partial L}{f_t} = \frac{\partial L}{\partial C_t} \frac{\partial C_t}{\partial f_t}$$

$$= \frac{\partial L}{\partial C_t} \frac{\partial}{\partial f_t} \left(C_{t-1} \otimes f_t + i_t \otimes \tilde{C}_t \right)$$

$$\Rightarrow \boxed{\frac{\partial L}{\partial f_t} = \frac{\partial L}{\partial C_t} C_{t-1}}$$

$$\frac{\partial L}{\partial a_f} = \frac{\partial L}{\partial f_t} \frac{\partial f_t}{\partial a_f}$$

$$\Rightarrow \boxed{\frac{\partial L}{\partial a_f} = \frac{\partial L}{\partial f_t} \left(\sigma(a_f) \left(1 - \sigma(a_f) \right) \right)}$$

$$\left\{ \begin{array}{ll}
\frac{\partial L}{\partial W_v} = \frac{\partial L}{\partial v_t} \frac{\partial v_t}{\partial W_v} = \frac{\partial L}{\partial v_t} h_t^T & \frac{\partial L}{\partial b_v} = \frac{\partial L}{\partial v_t} \frac{\partial v_t}{\partial b_v} = \frac{\partial L}{\partial v_t} \\
\frac{\partial L}{\partial W_o} = \frac{\partial L}{\partial o_t} \frac{\partial o_t}{\partial W_o} = \frac{\partial L}{\partial o_t} z_t^T & , \frac{\partial L}{\partial b_v} = \frac{\partial L}{\partial o_t} \frac{\partial o_t}{\partial b_o} = \frac{\partial L}{\partial o_t} \\
\frac{\partial L}{\partial W_{\tilde{c}}} = \frac{\partial L}{\partial \tilde{C}_t} \frac{\partial \tilde{C}}{\partial W_{\tilde{c}}} = \frac{\partial L}{\partial \tilde{C}_t} z_t^T & , \frac{\partial L}{\partial b_{\tilde{c}}} = \frac{\partial L}{\partial \tilde{C}_t} \frac{\partial \tilde{C}_t}{\partial b_o} = \frac{\partial L}{\partial \tilde{C}_t} \\
\frac{\partial L}{\partial W_i} = \frac{\partial L}{\partial i_t} \frac{\partial i_t}{\partial W_i} = \frac{\partial L}{\partial i_t} z_t^T & \frac{\partial L}{\partial b_i} = \frac{\partial L}{\partial i_t} \frac{\partial i_t}{\partial b_i} = \frac{\partial L}{\partial i_t} \\
\frac{\partial L}{\partial W_f} = \frac{\partial L}{\partial f_t} \frac{\partial f_t}{\partial W_f} = \frac{\partial L}{\partial f_t} z_t^T & \frac{\partial L}{\partial b_f} = \frac{\partial L}{\partial f_t} \frac{\partial f_t}{\partial b_f} = \frac{\partial L}{\partial f_t}
\end{array} \right.$$

$$\frac{\partial L}{\partial z_t} = \frac{\partial L}{\partial a_f} \frac{\partial a_f}{\partial z_t} + \frac{\partial L}{\partial a_i} \frac{\partial a_i}{\partial z_t} + \frac{\partial L}{\partial a_{\tilde{c}}} \frac{\partial a_{\tilde{c}}}{\partial z_t} + \frac{\partial L}{\partial a_o} \frac{\partial a_o}{\partial z_t}$$

$$= \frac{\partial L}{\partial a_f} W_a^T + \frac{\partial L}{\partial a_i} W_i^T + \frac{\partial L}{\partial a_{\tilde{c}}} W_{\tilde{c}}^T + \frac{\partial L}{\partial a_o} W_o^T$$