

LSTM formula

$$\begin{bmatrix} i \\ f \\ o \\ g \end{bmatrix} = \begin{bmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{bmatrix} W \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix}$$

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

$$= \begin{bmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{bmatrix} \begin{bmatrix} W_i \\ W_f \\ W_o \\ W_g \end{bmatrix} \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix}$$

$$= \begin{bmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{bmatrix} \begin{bmatrix} W_{hi} h_{t-1} + W_{xi} x_t \\ W_{hf} h_{t-1} + W_{xf} x_t \\ W_{ho} h_{t-1} + W_{xo} x_t \\ W_{hg} h_{t-1} + W_{xg} x_t \end{bmatrix}$$

$$C_t = f \odot C_{t-1} + i \odot g$$

$$h_t = o \odot \tanh(C_t)$$

RNN formula

$$h_t = \tanh \left[ W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} \right]$$

$$= \tanh [W_{hh} h_{t-1} + W_{xh} x_t]$$