

RNN



فصل پنجم



شرکت فروش مصنوعي رسا

مدرس : خريد فاشمي نژاد

LSTM & GRU



جلسہ پیست و فگتم



شرکت فروش مصنوعی رسا

مدرس : خرید فاشمی نژاد

- راه حل های بسیاری برای Vanishing Gradients مطرح شد

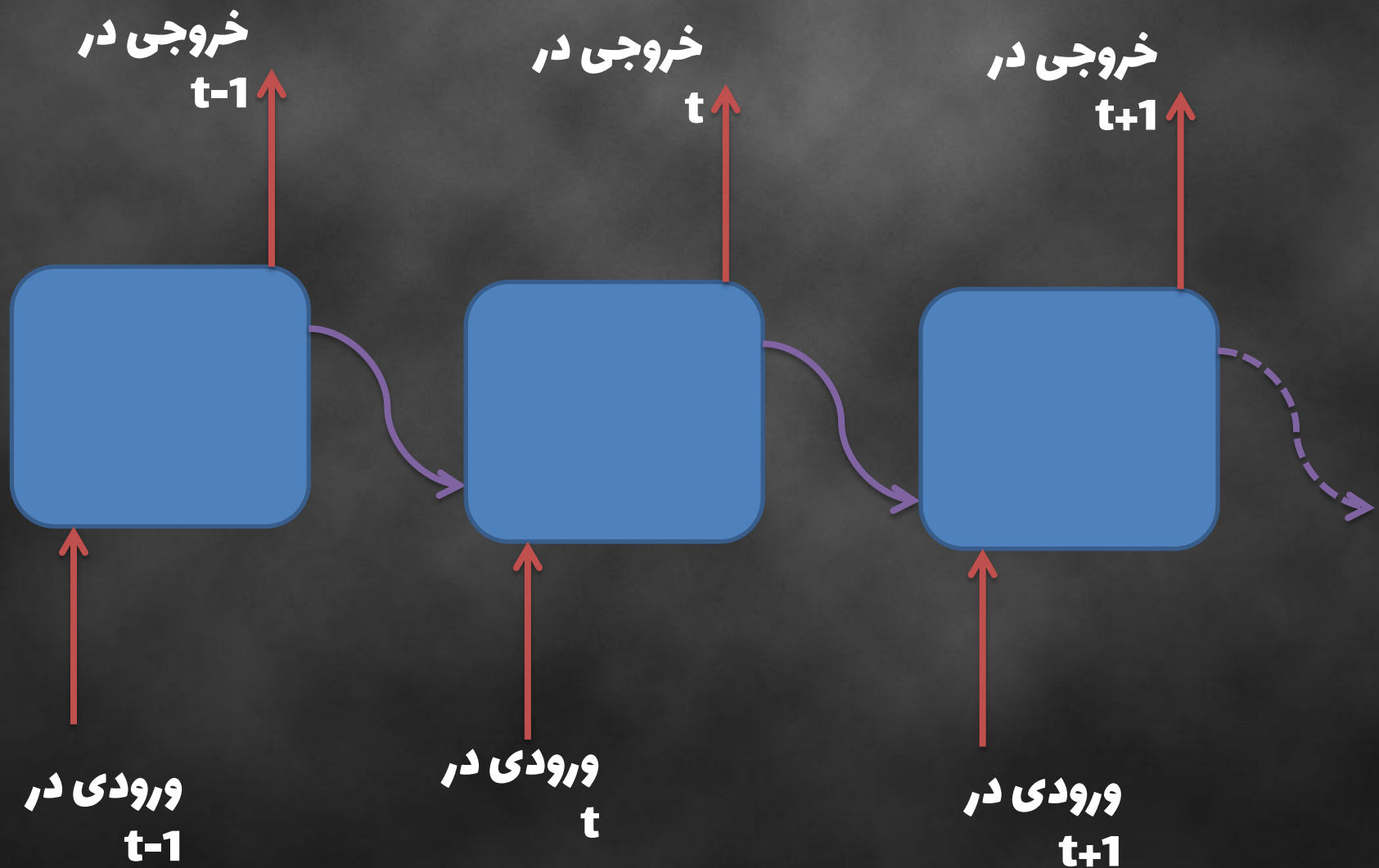
- ولی وقتی طول سری زمانی بالا باشد روند آموزش زمان بر خواهد شد

- راه اولیه: کار کردن با RNN در سری زمانی های کوتاه است؟

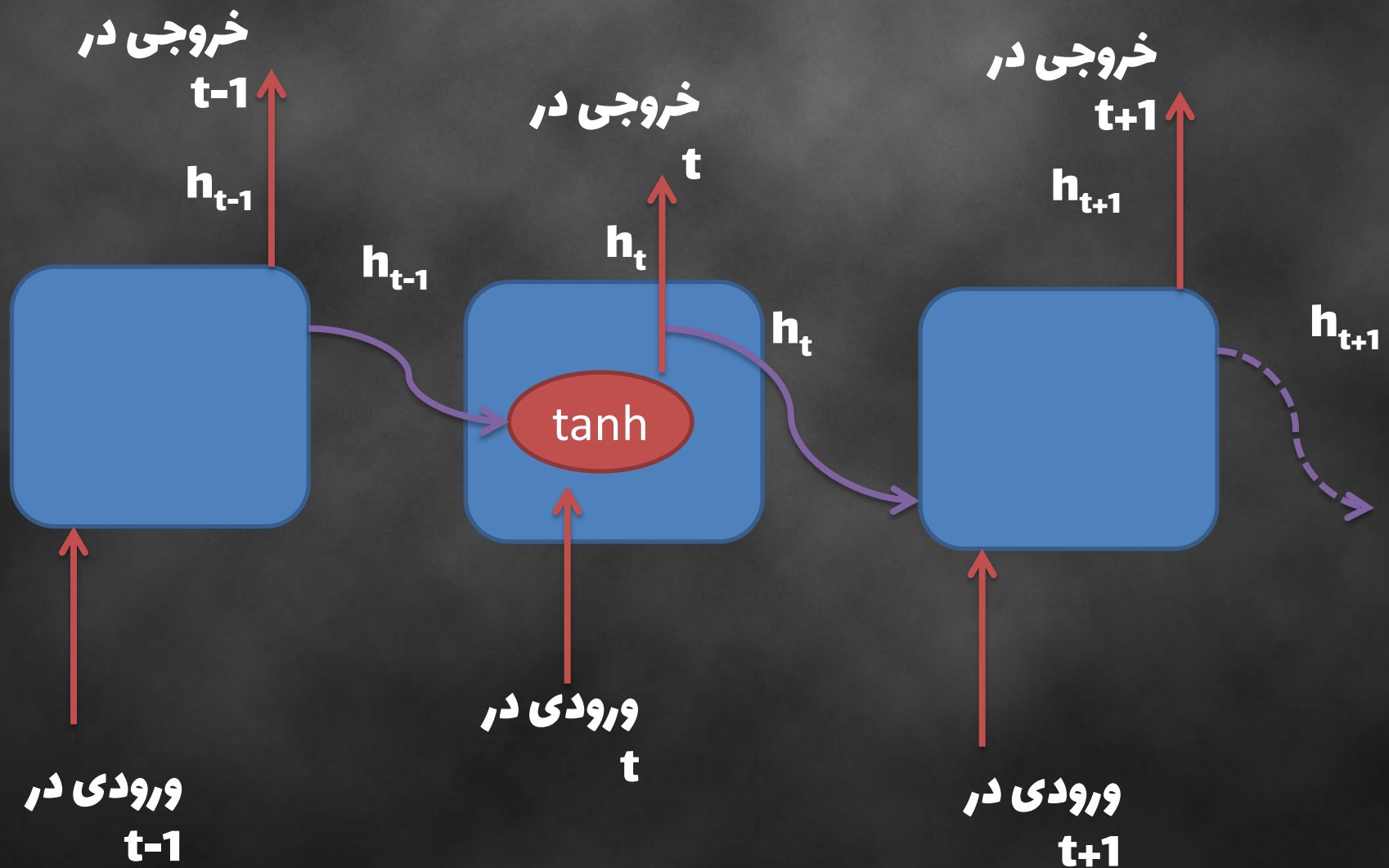
- آخر سری زمانی بلند داشتیم چه؟

- LSTM
- Long Short Term Memory
- چطوری کار میکنه؟

RNN معمولی



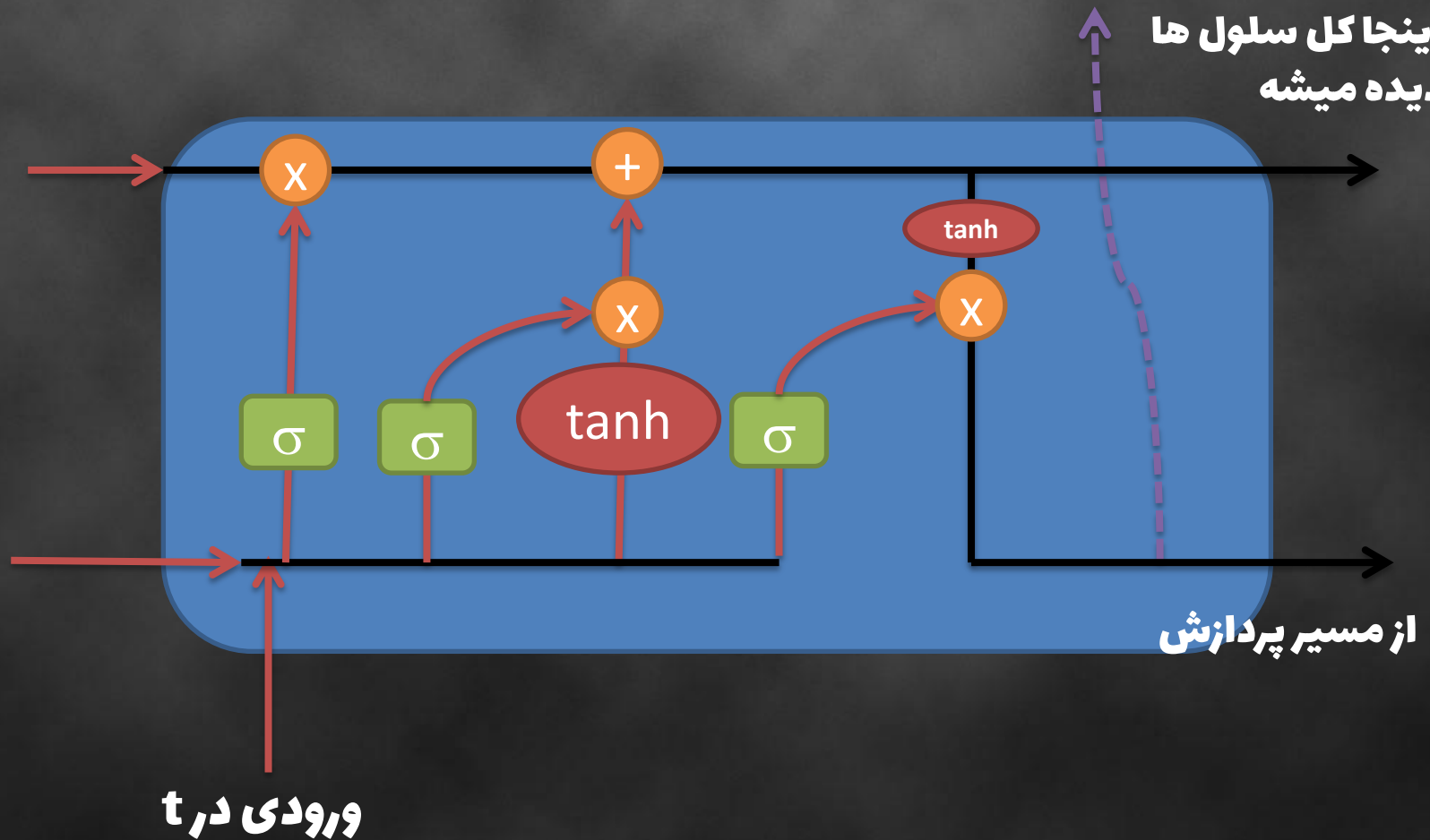
RNN معمولی



• سلول LSTM

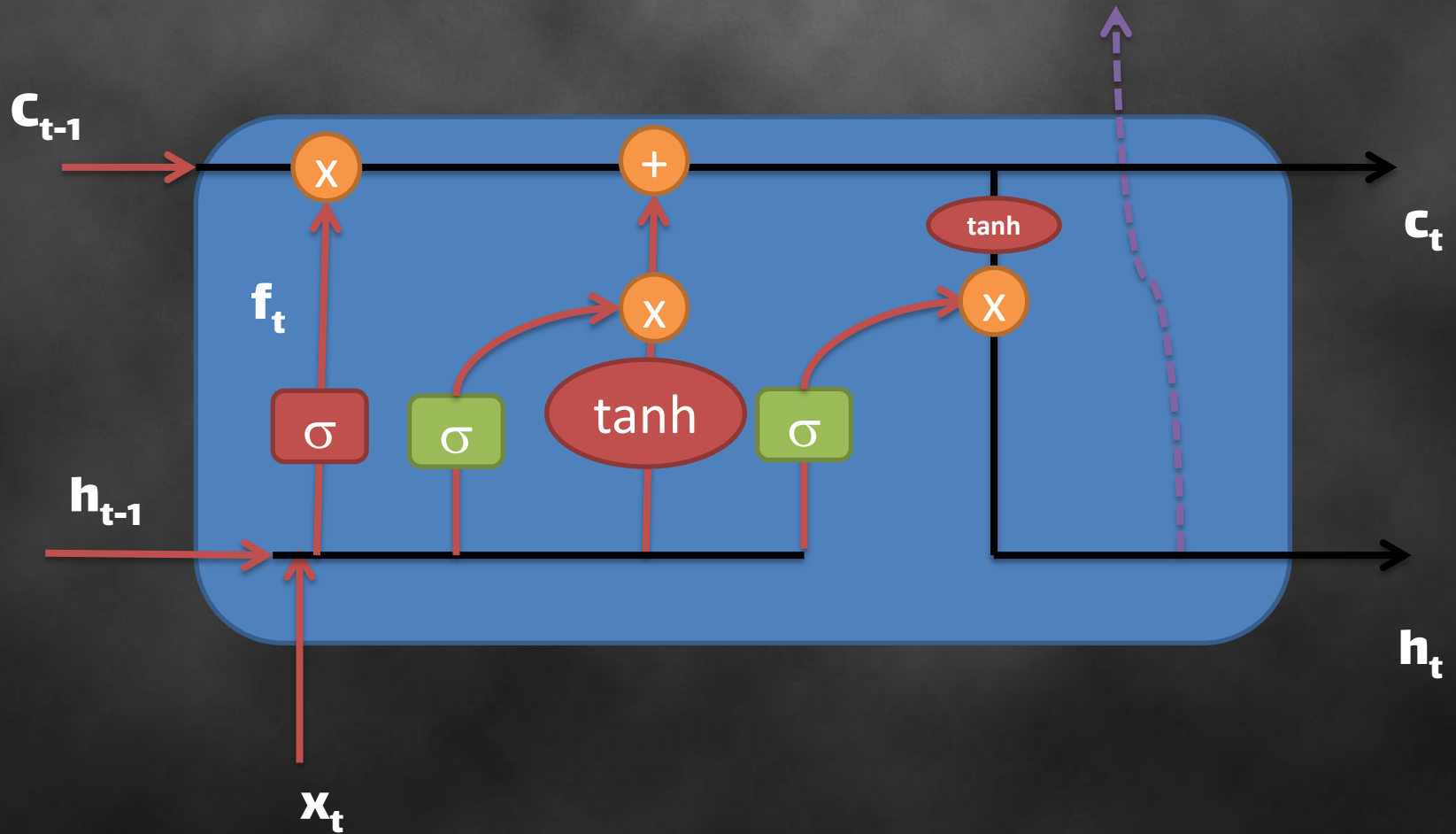
خروجی در t

اینجا کل سلول ها
دیده میشه



LSTM سلول •

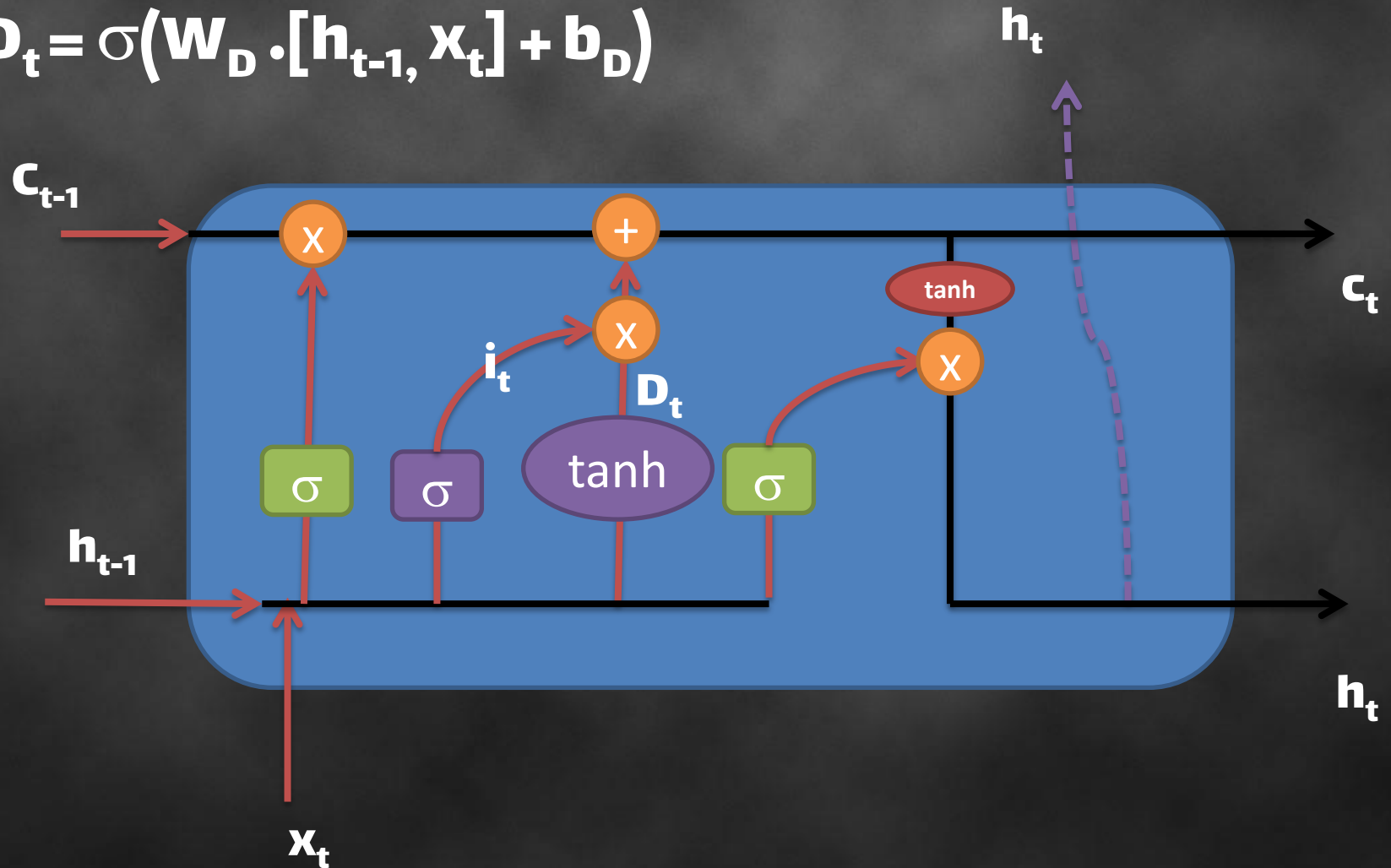
$$\mathbf{f}_t = \sigma(\mathbf{w}_f \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f)$$



$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i)$$

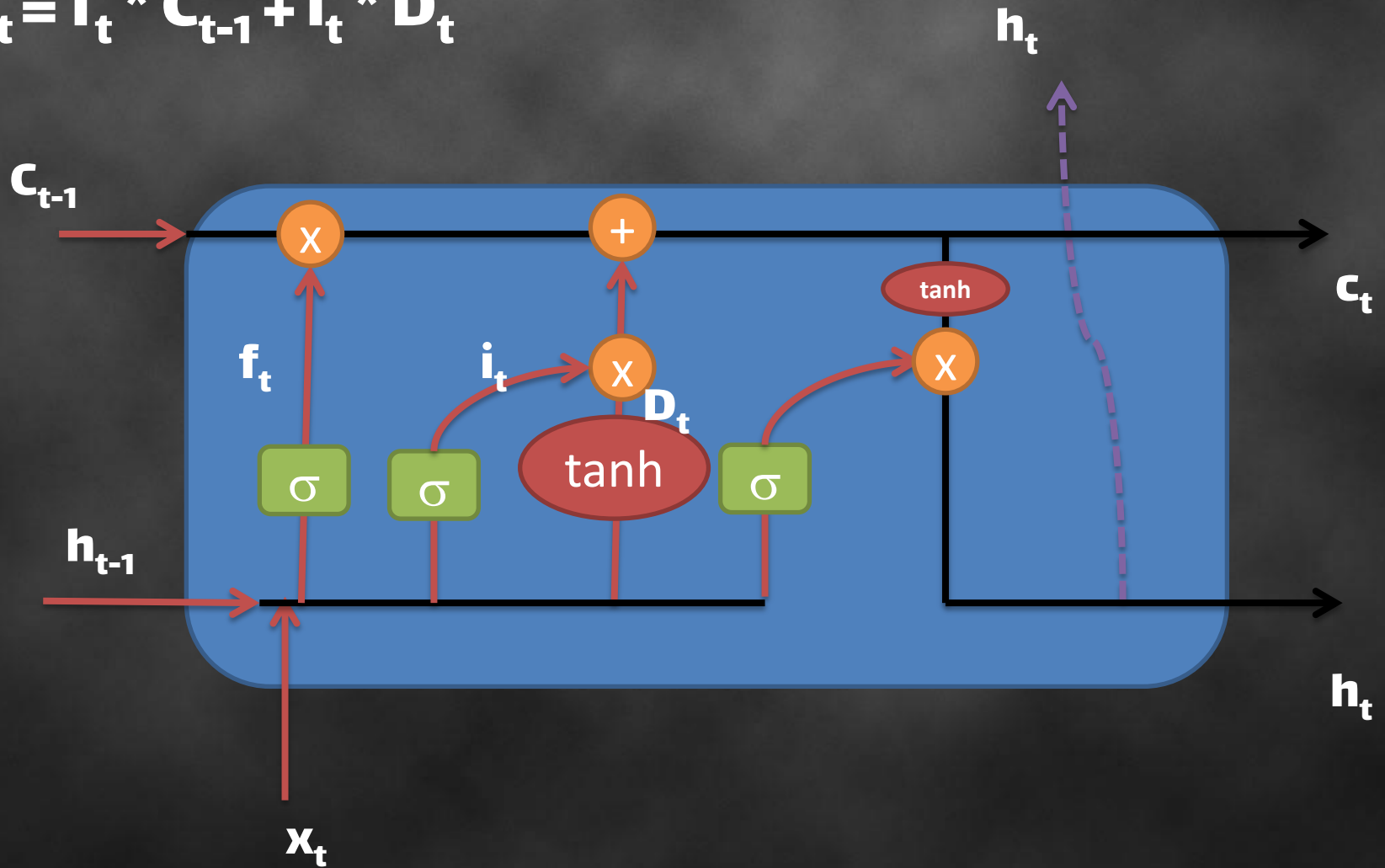
$$D_t = \sigma(w_D \cdot [h_{t-1}, x_t] + b_D)$$

LSTM سلول •



LSTM سلول •

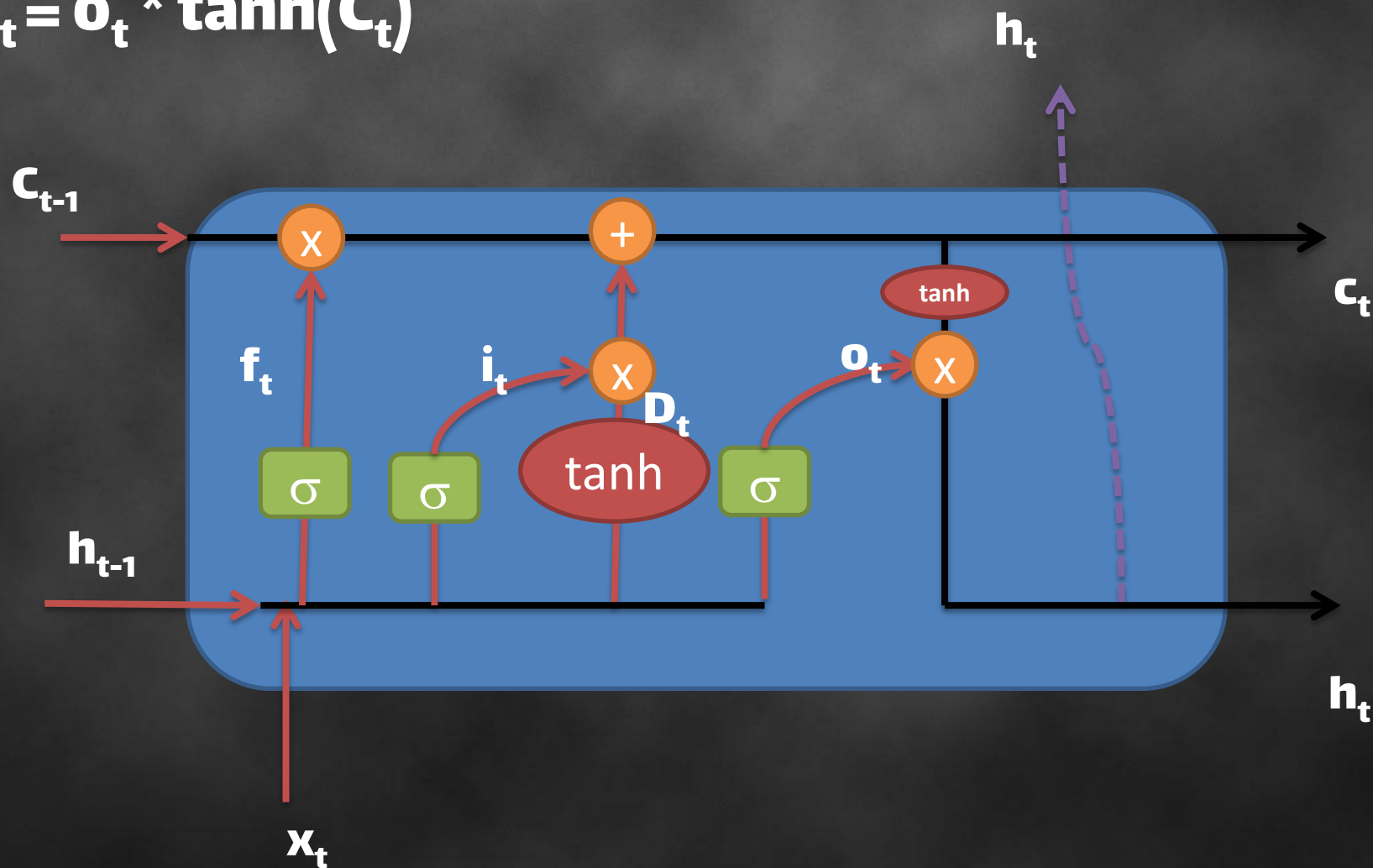
$$C_t = f_t * C_{t-1} + i_t * D_t$$



LSTM سلول •

$$\mathbf{o}_t = \sigma(\mathbf{W}_o \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o)$$

$$\mathbf{h}_t = \mathbf{o}_t * \tanh(\mathbf{C}_t)$$

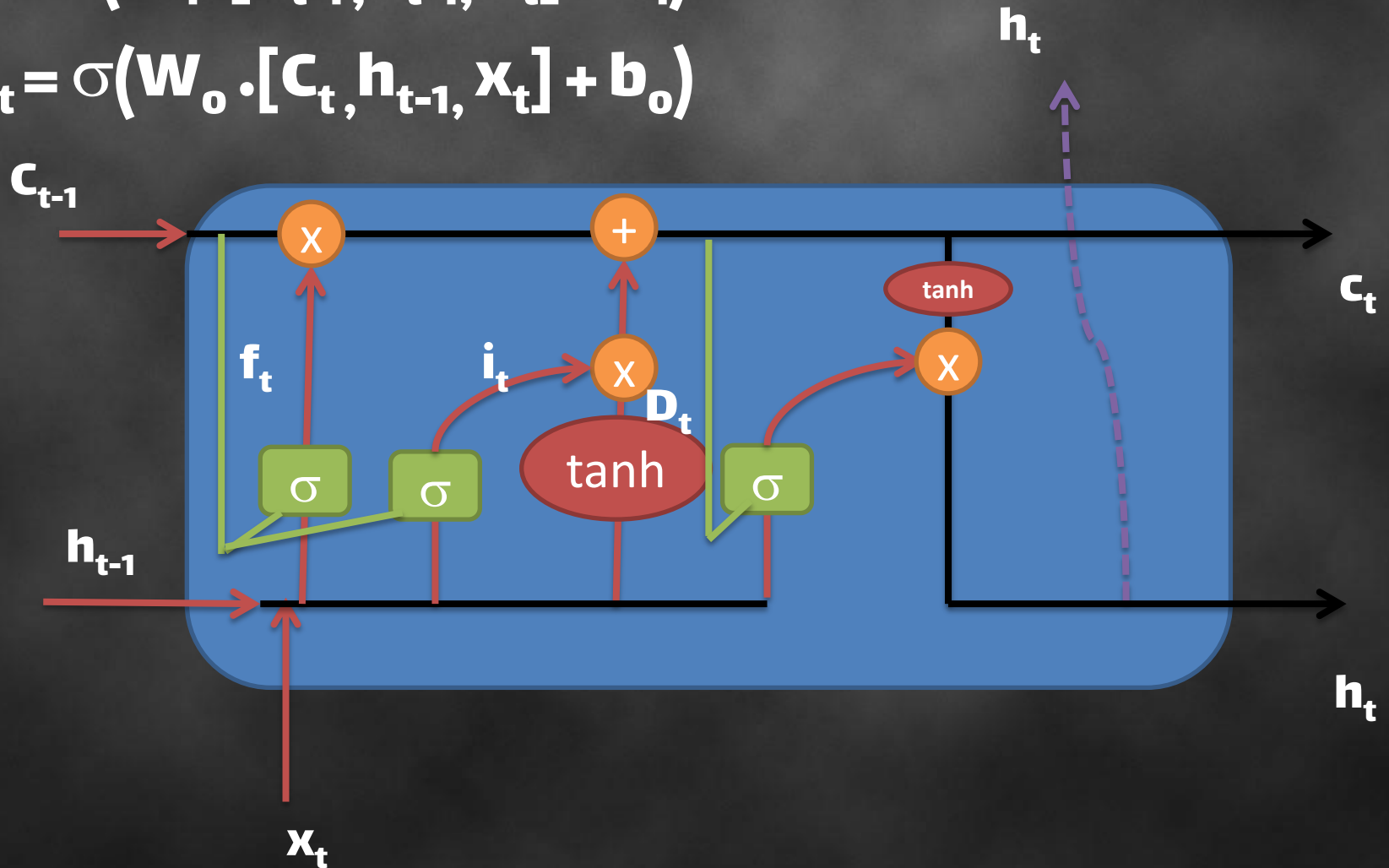


$$\mathbf{f}_t = \sigma(\mathbf{W}_f \cdot [\mathbf{C}_{t-1}, \mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f)$$

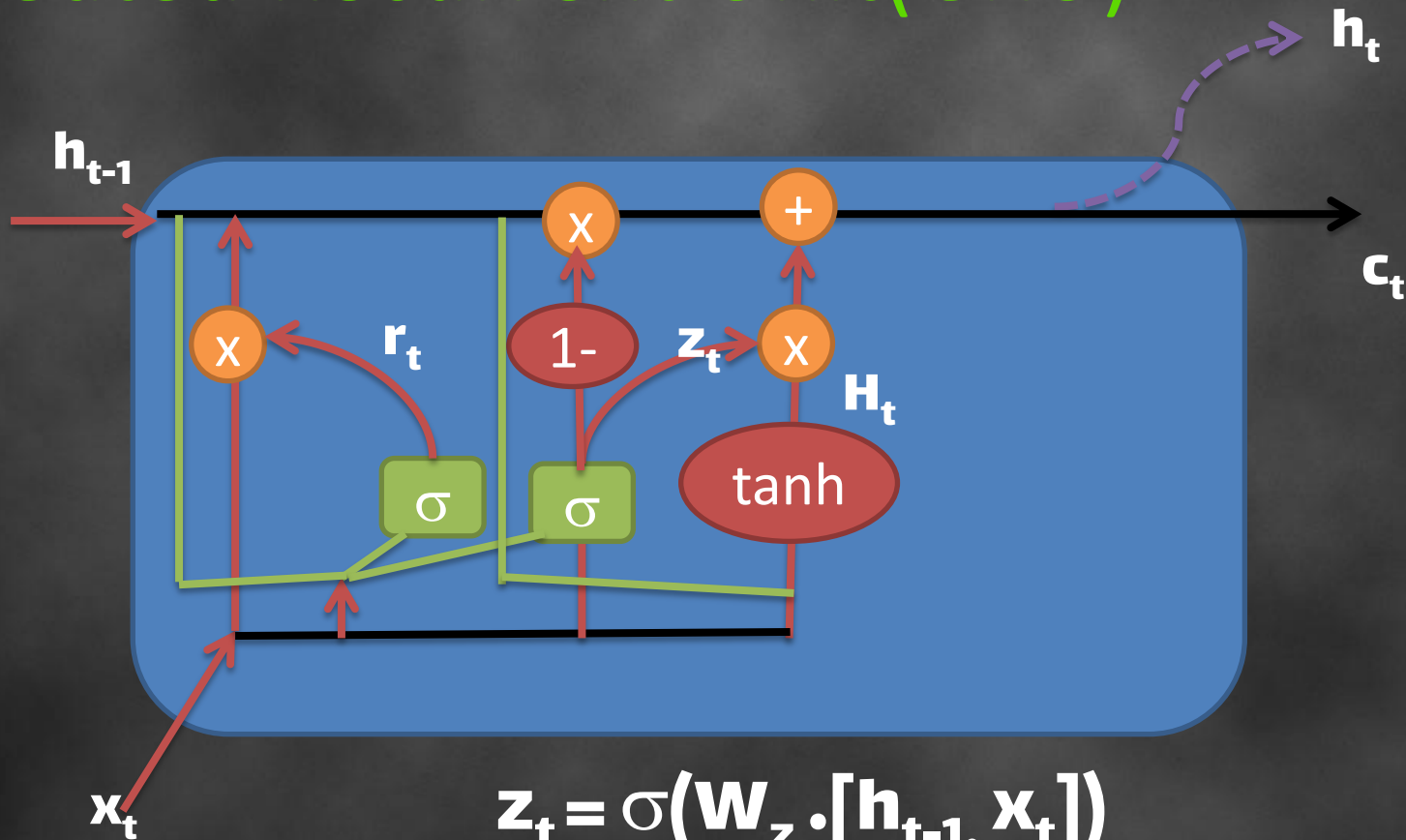
$$\mathbf{i}_t = \sigma(\mathbf{W}_i \cdot [\mathbf{C}_{t-1}, \mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i)$$

$$\mathbf{o}_t = \sigma(\mathbf{W}_o \cdot [\mathbf{C}_t, \mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o)$$

LSTM سلول •



Gated Recurrent Unit (GRU)



$$z_t = \sigma(w_z \cdot [h_{t-1}, x_t])$$

$$r_t = \sigma(w_r \cdot [h_{t-1}, x_t])$$

$$H_t = \tanh(w_H \cdot [r_t * h_{t-1}, x_t])$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * H_t$$