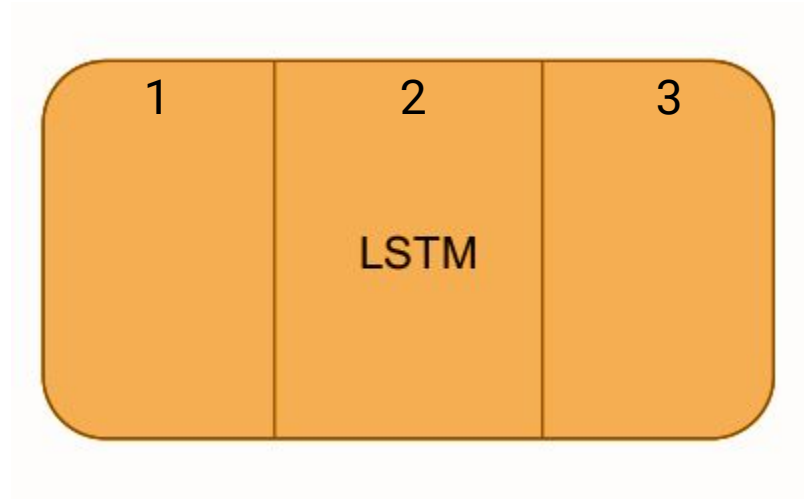


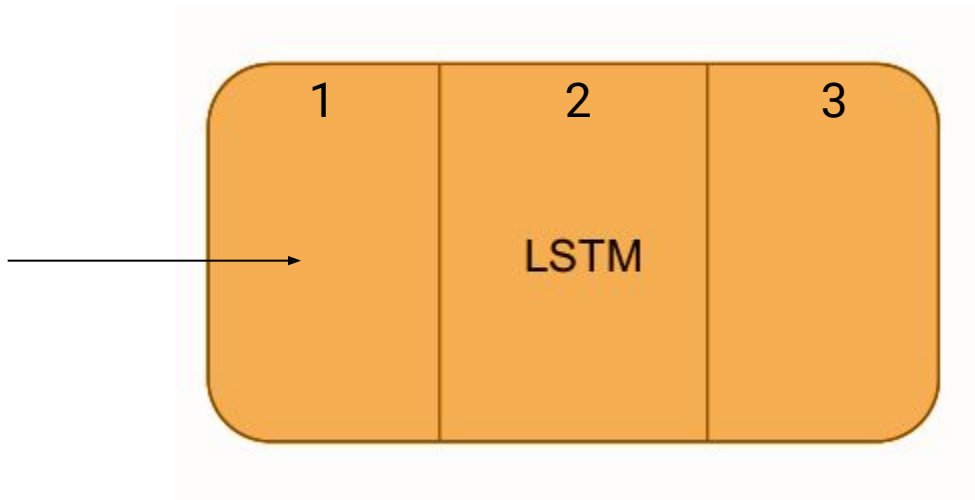
Long Short Term Memory (LSTM)

LSTM



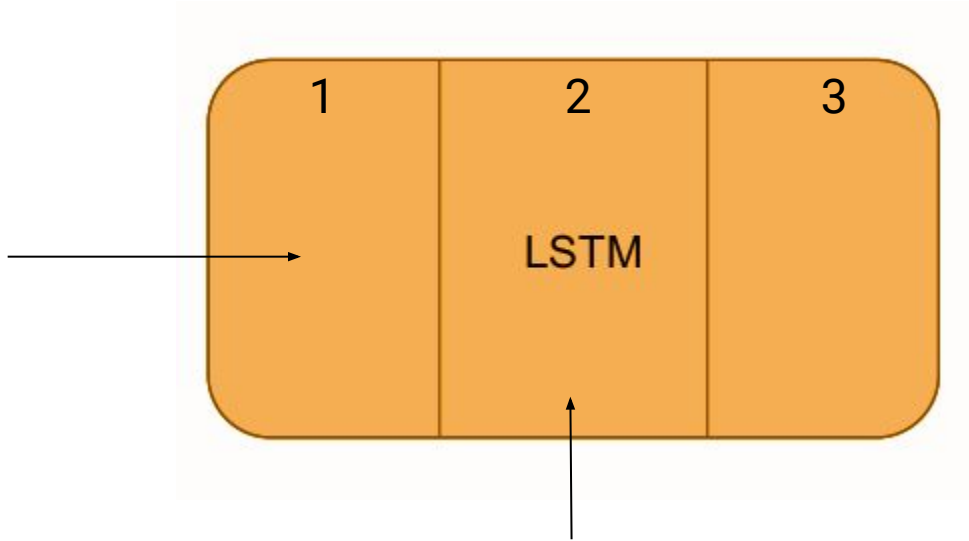
LSTM

Forget irrelevant
information



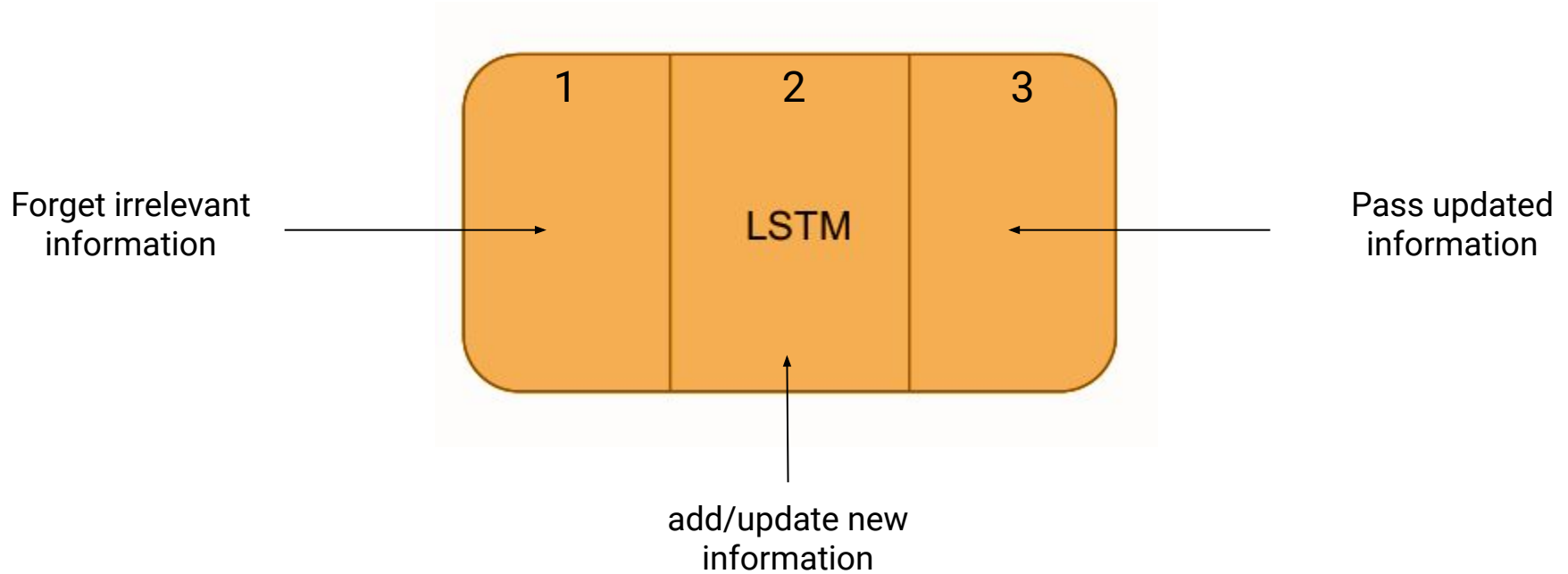
LSTM

Forget irrelevant
information

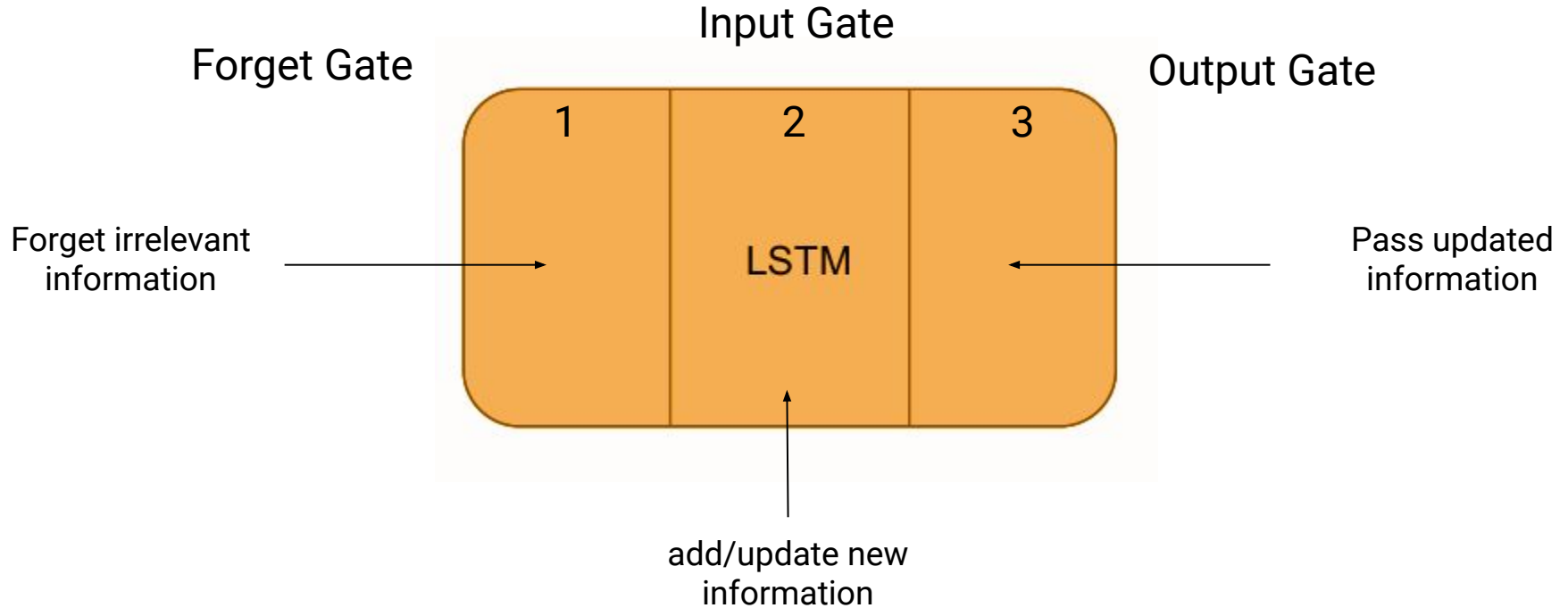


add/update new
information

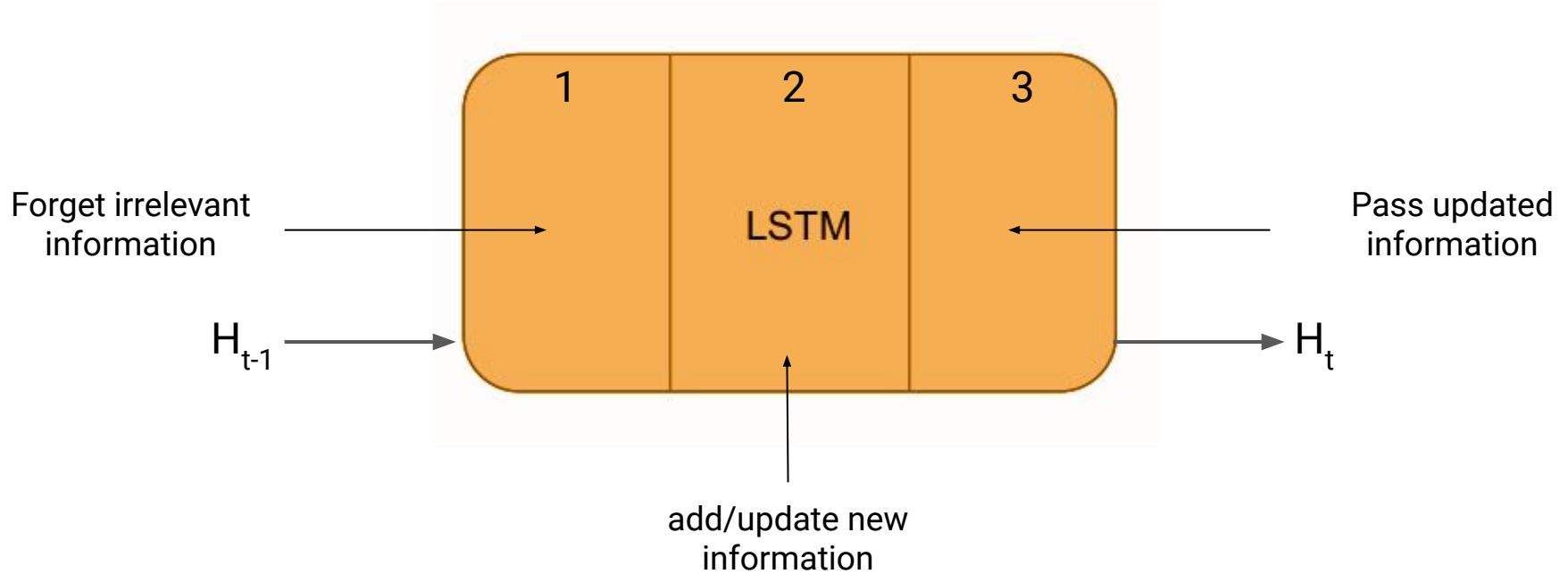
LSTM



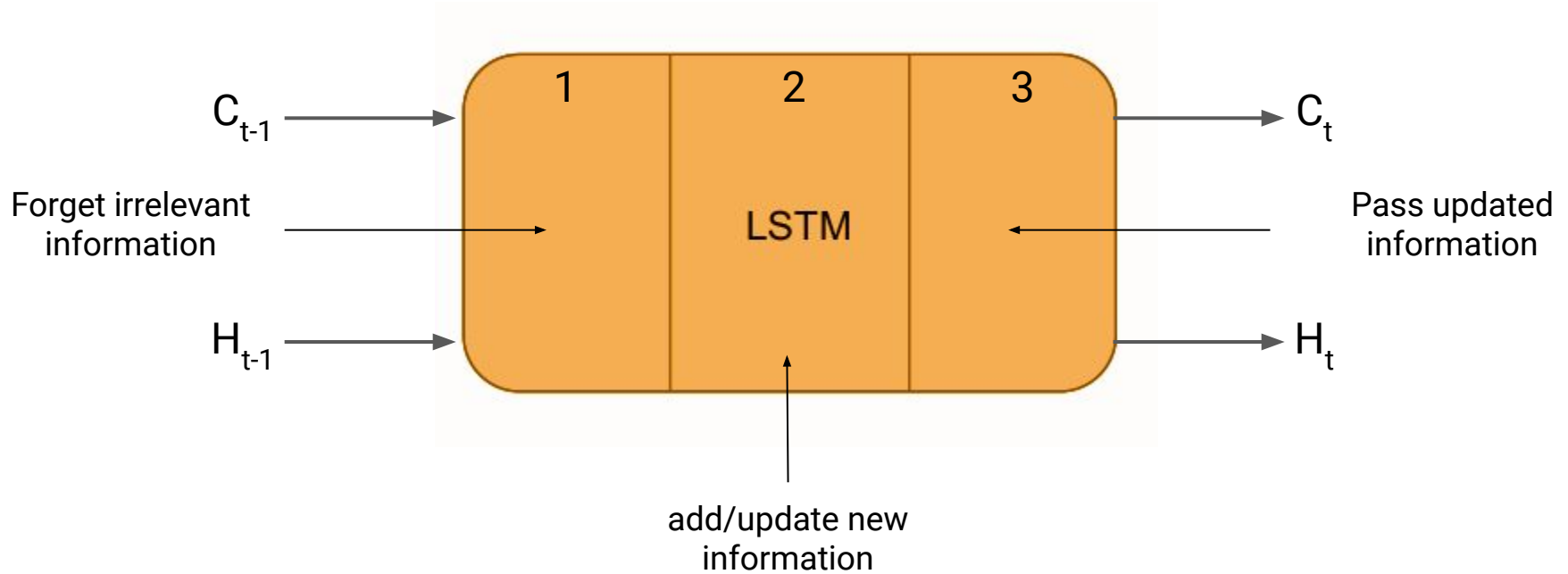
LSTM



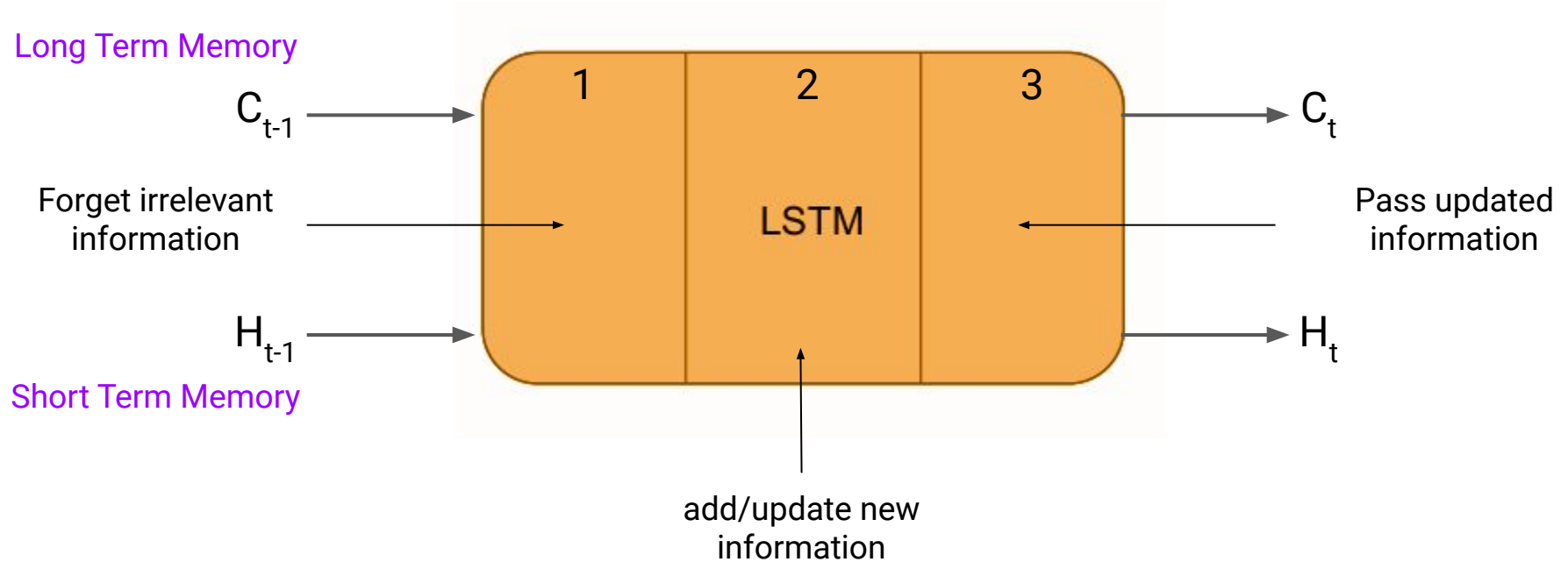
LSTM



LSTM



LSTM



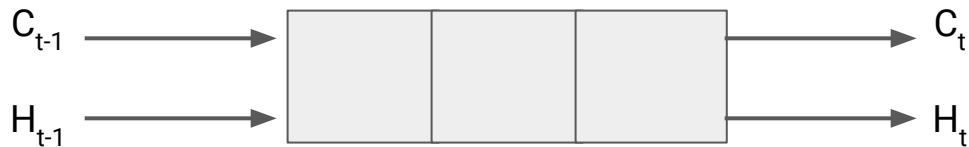
LSTM



LSTM

Bob is a nice person. Dan on the other hand is evil.

LSTM



LSTM

Bob is a nice person. Dan on the other hand is evil.

LSTM



LSTM

Bob is a nice person. Dan on the other hand is evil.

Forget Gate

LSTM

Forget Gate:

- $f_t = \sigma(x_t * U_f + H_{t-1} * W_f)$

LSTM

Forget Gate:

- $f_t = \sigma(x_t * U_f + H_{t-1} * W_f)$

$$C_{t-1} * f_t = 0 \quad \dots \text{if } f_t = 0 \text{ (forget everything)}$$

LSTM

Forget Gate:

- $f_t = \sigma(x_t * U_f + H_{t-1} * W_f)$

$$C_{t-1} * f_t = 0 \quad \dots \text{if } f_t = 0 \text{ (forget everything)}$$

$$C_{t-1} * f_t = C_{t-1} \quad \dots \text{if } f_t = 1 \text{ (forget nothing)}$$

LSTM

Forget Gate:

- $f_t = \sigma(x_t * U_f + H_{t-1} * W_f)$

$$C_{t-1} * f_t$$

Bob is a nice person. Dan on the other hand is evil.

LSTM

Forget Gate:

- $f_t = \sigma(x_t * U_f + H_{t-1} * W_f)$

$$C_{t-1} * f_t$$

Bob is a nice person. Dan on the other hand is evil.

Bob is a nice person. Dan on the other hand is evil.

LSTM



LSTM

Bob knows swimming. He told me over the phone that he had served the navy for 4 long years.

LSTM



LSTM

Bob knows **swimming**. He told me over the **phone** that he had served the **navy** for 4 long years.

LSTM



LSTM

Bob knows **swimming**. He told me over the **phone** that he had served the **navy** for 4 long years.

Input Gate

LSTM

Input Gate:

- $i_t = \sigma(x_t * U_i + H_{t-1} * W_i)$

He told me over the phone...
served in the navy

LSTM

Input Gate:

- $i_t = \sigma(x_t * U_i + H_{t-1} * W_i)$
- $N_t = \tanh(x_t * U_c + H_{t-1} * W_c)$ (new information)

He told me over the phone...
served in the navy

LSTM

Input Gate:

- $i_t = \sigma(x_t * U_i + H_{t-1} * W_i)$
- $N_t = \tanh(x_t * U_c + H_{t-1} * W_c)$ (new information)
- $C_t = f_t * C_{t-1} + i_t * N_t$ (updating cell state)

He told me over the phone...
served in the navy

LSTM



LSTM

Bob single handedly fought the enemy and died for his country. For his contributions, brave _____.

LSTM



LSTM

Bob single handedly fought the enemy and died for his country. For his contributions, brave _____.

LSTM



LSTM

Bob single handedly fought the enemy and died for his country. For his contributions, **brave** _____.

Output Gate

LSTM

Output Gate:

- $$o_t = \sigma (x_t * U_o + H_{t-1} * W_o)$$

LSTM

Output Gate:

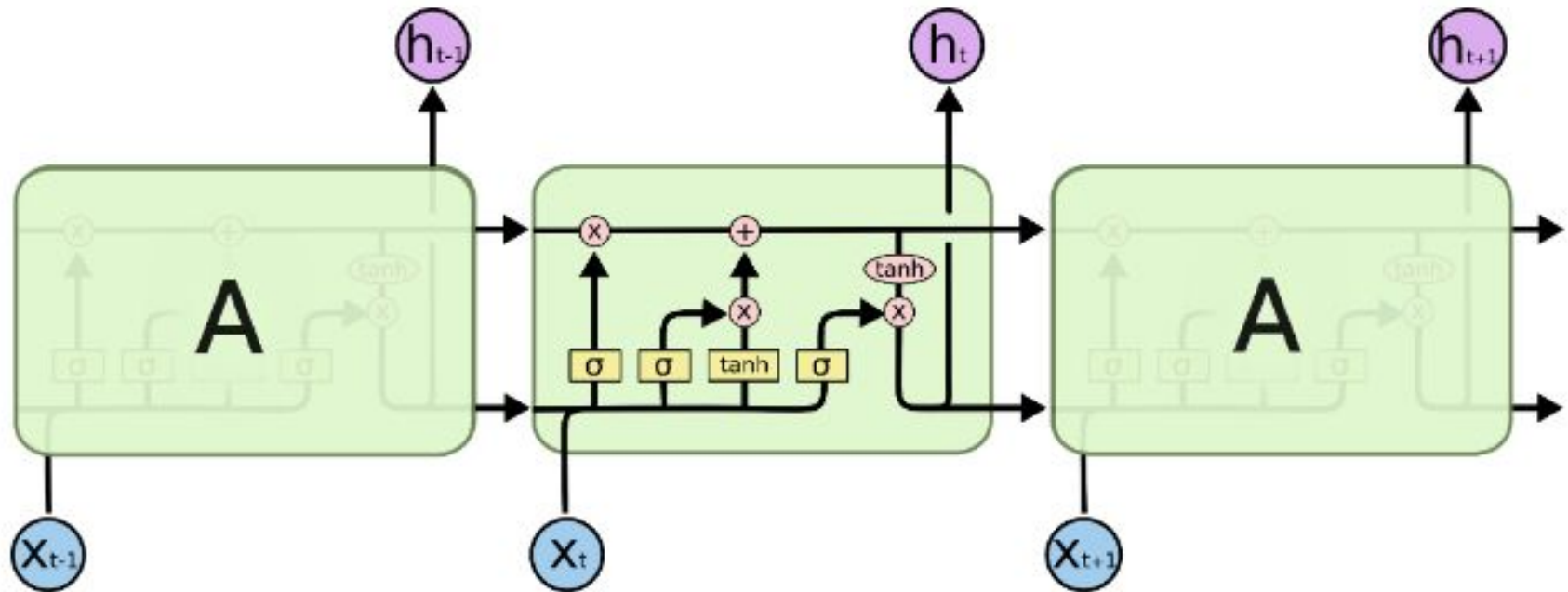
- $o_t = \sigma(x_t * U_o + H_{t-1} * W_o)$
- $H_t = o_t * \tanh(C_t)$

LSTM

Output Gate:

- $o_t = \sigma (x_t * U_o + H_{t-1} * W_o)$
- $H_t = o_t * \tanh(C_t)$
- $\text{Output} = \text{Softmax}(H_t)$

LSTM



Thank You