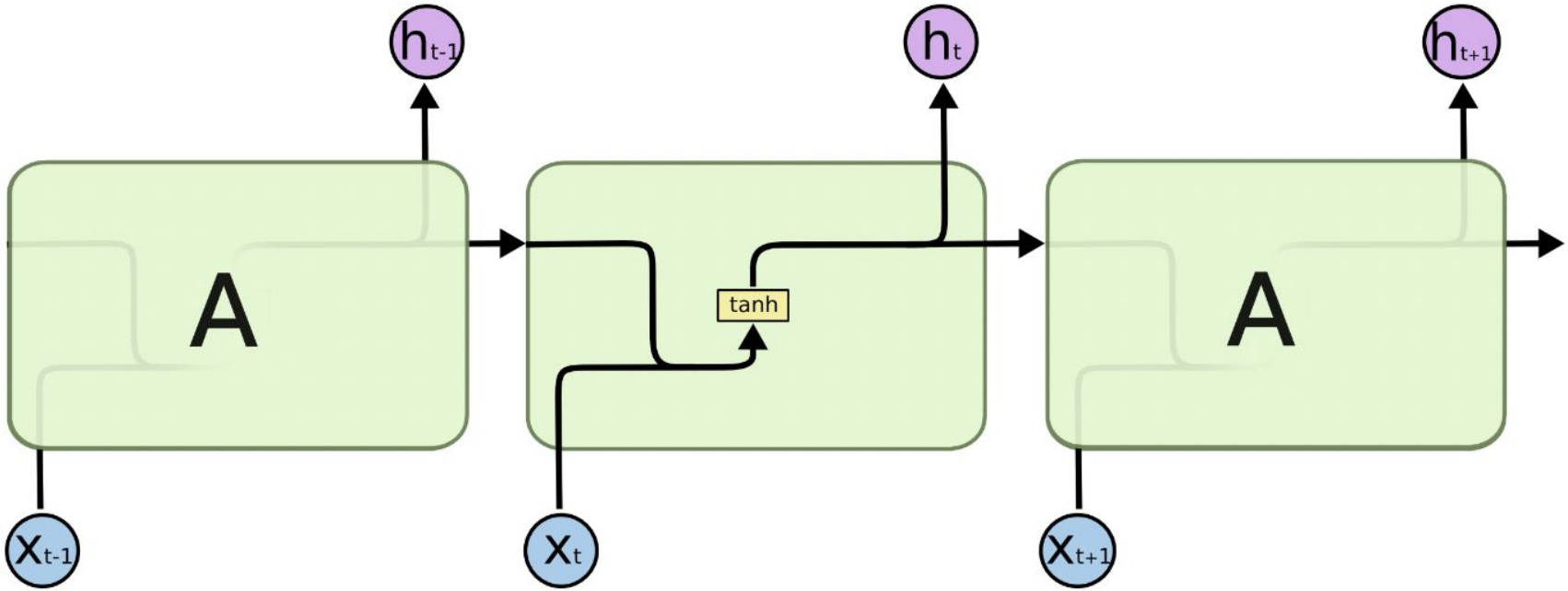
Advanced Machine Learning

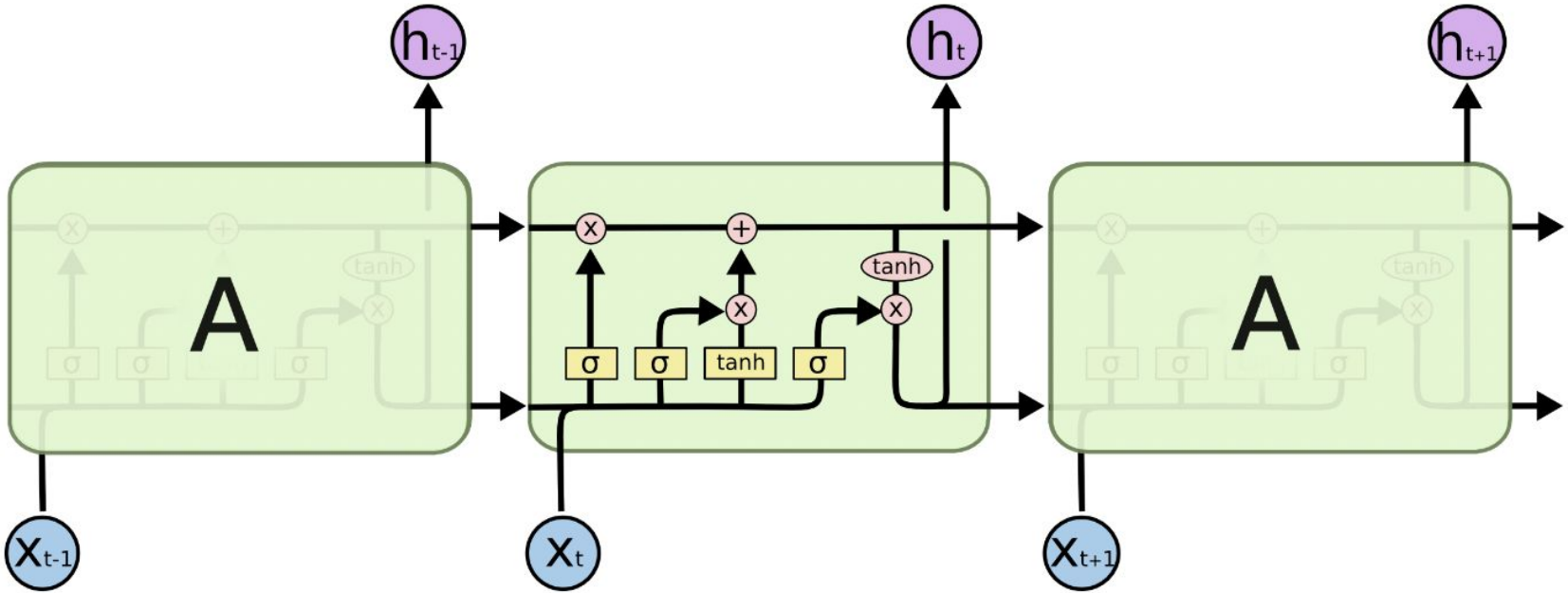
Likhit Nayak

Long short-term memory (LSTM)



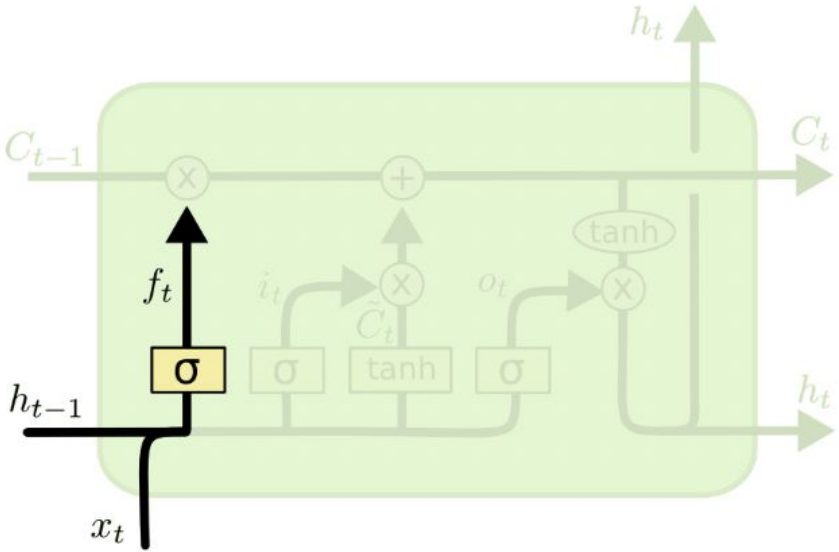
Understanding LSTM Networks -- colah's blog. (n.d.). https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Long short-term memory



Understanding LSTM Networks -- colah's blog. (n.d.). https://colah.github.io/posts/2015-08-Understanding-LSTMs/

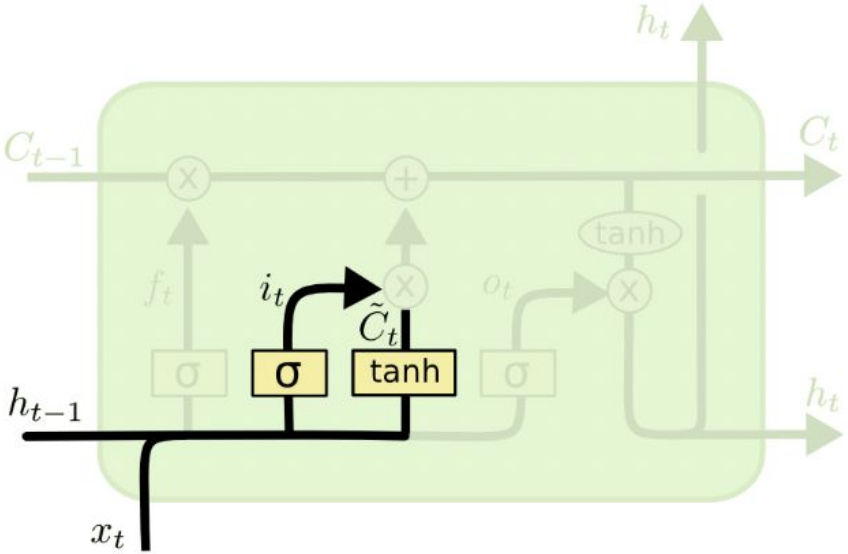
Long short-term memory



$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \; + \; b_f\right)$$

Understanding LSTM Networks -- colah's blog. (n.d.). https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Long short-term memory

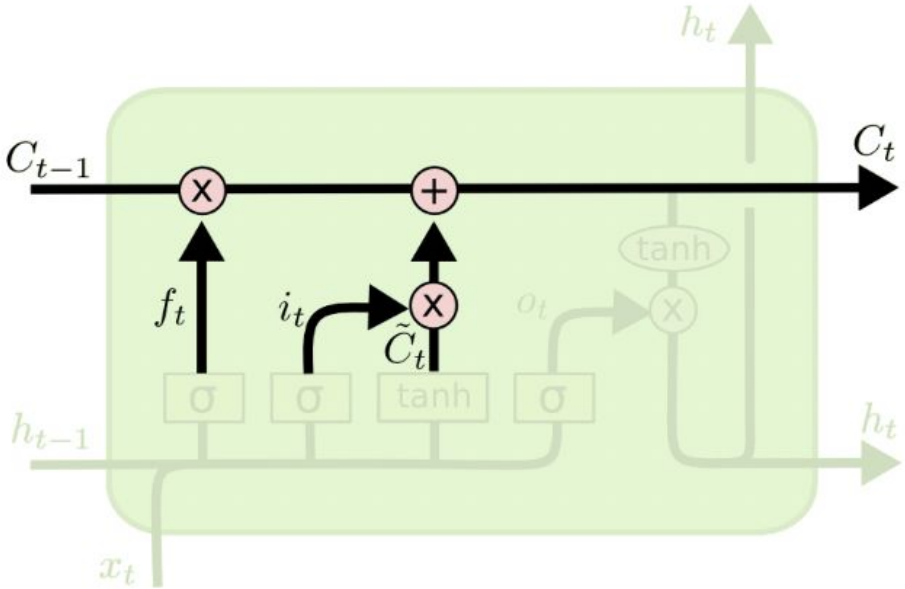


$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \; + \; b_i\right)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \; + \; b_C)$$

Understanding LSTM Networks -- colah's blog. (n.d.). https://colah.github.io/posts/2015-08-Understanding-LSTMs/

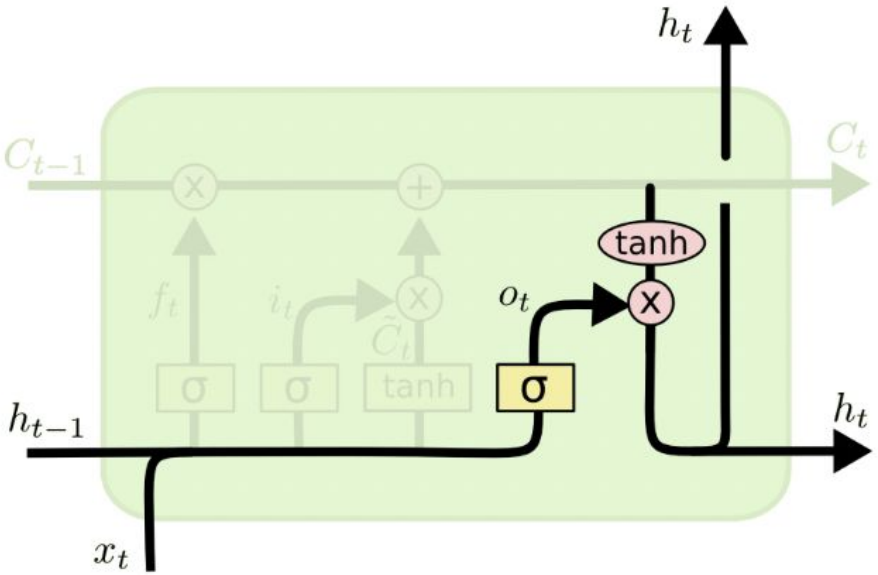
Long short-term memory



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

Understanding LSTM Networks -- colah's blog. (n.d.). https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Long short-term memory

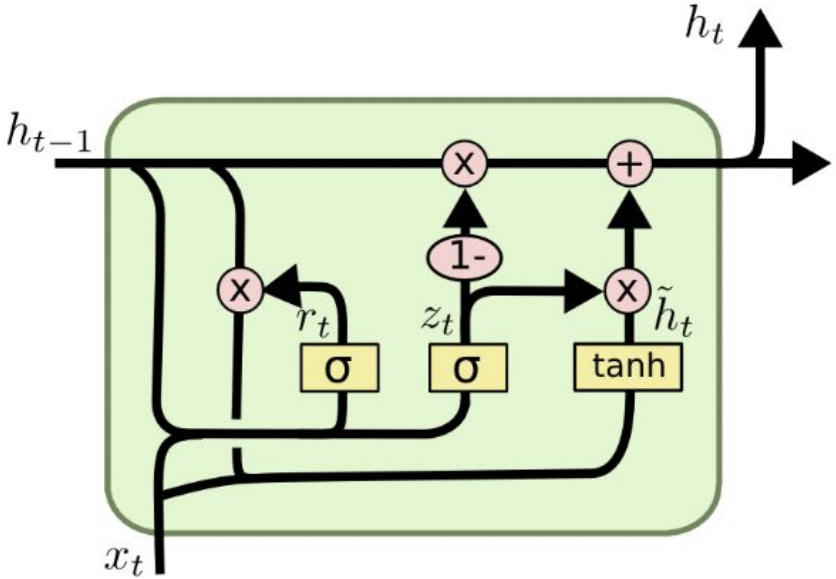


$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right)$$

$$h_t = o_t * \tanh\left(C_t\right)$$

Understanding LSTM Networks -- colah's blog. (n.d.). https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Gated Recurrent Unit (GRU)



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

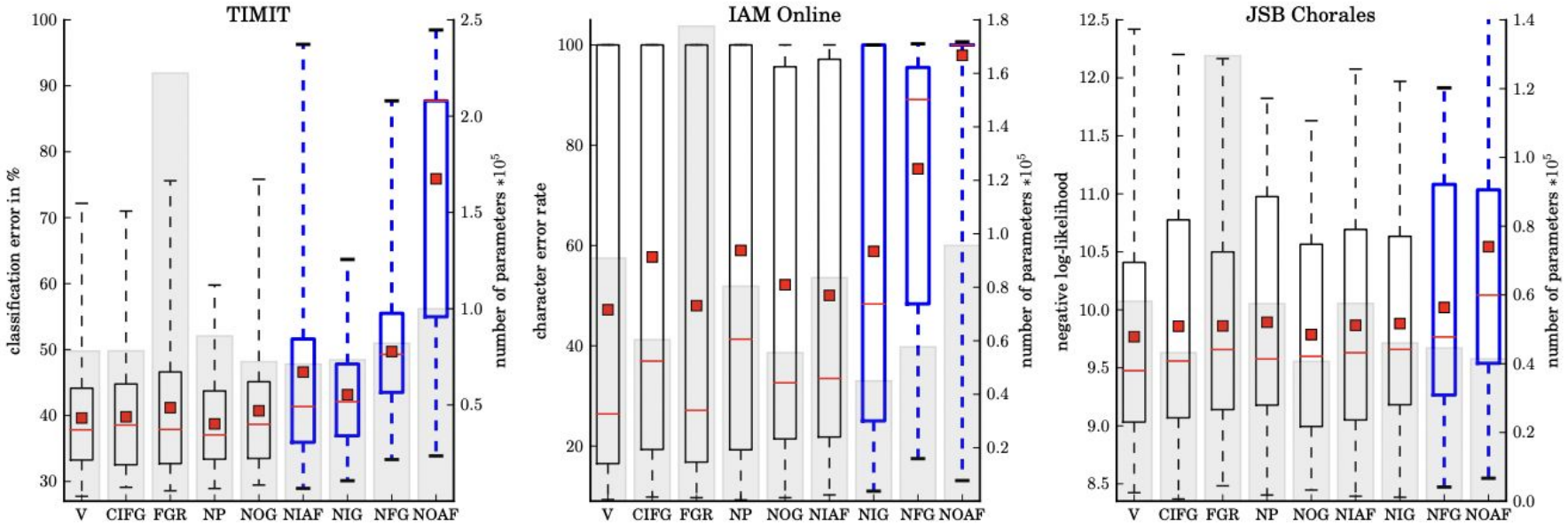
$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right)$$

$$\tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

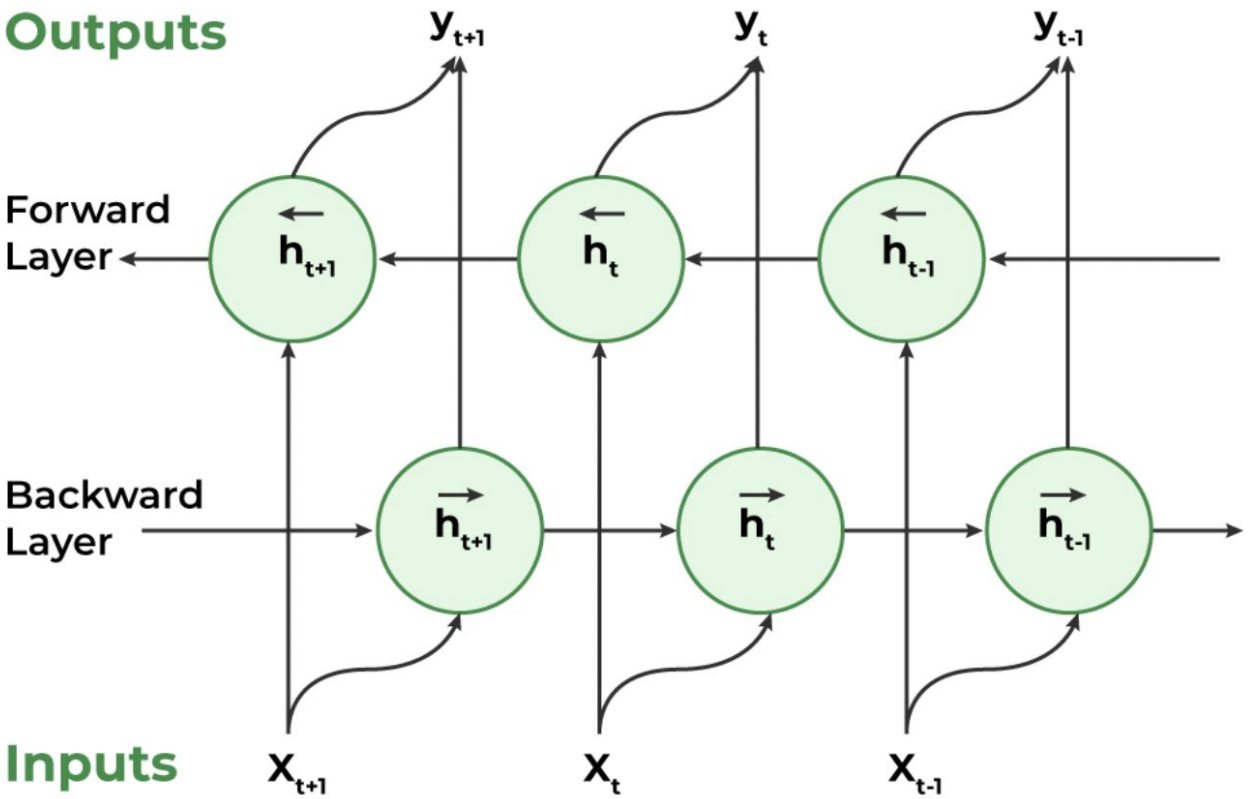
Understanding LSTM Networks -- colah's blog. (n.d.). https://colah.github.io/posts/2015-08-Understanding-LSTMs/

LSTM and its variants



Greff, Klaus, et al. "LSTM: A search space odyssey." IEEE transactions on neural networks and learning systems 28.10 (2016): 2222-2232.

Bidirectional RNNs



Deep RNNs