

# DEEP LEARNING – Recurrent Neural Network

Trainer: Dr. Darshan Ingle.



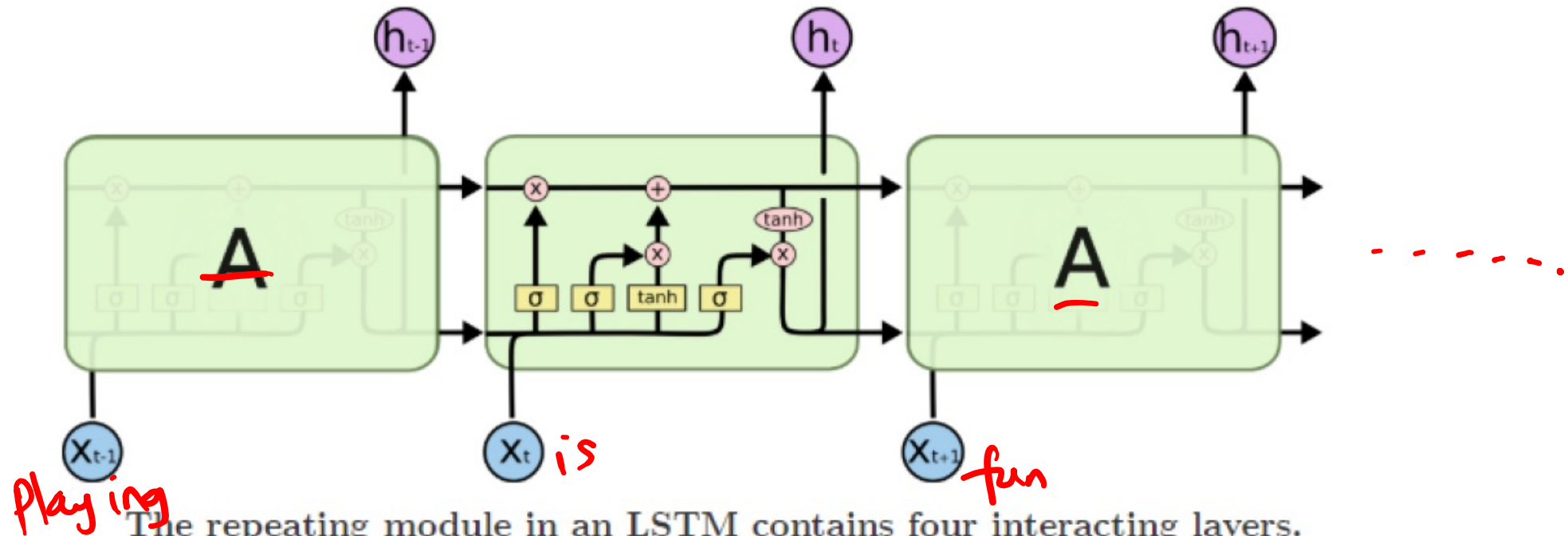
Problems with RNN are good in handling sequential data, but they run into a problem when the context is far away.

I live in Gujarat & I know Gujrati. [Context - Near]

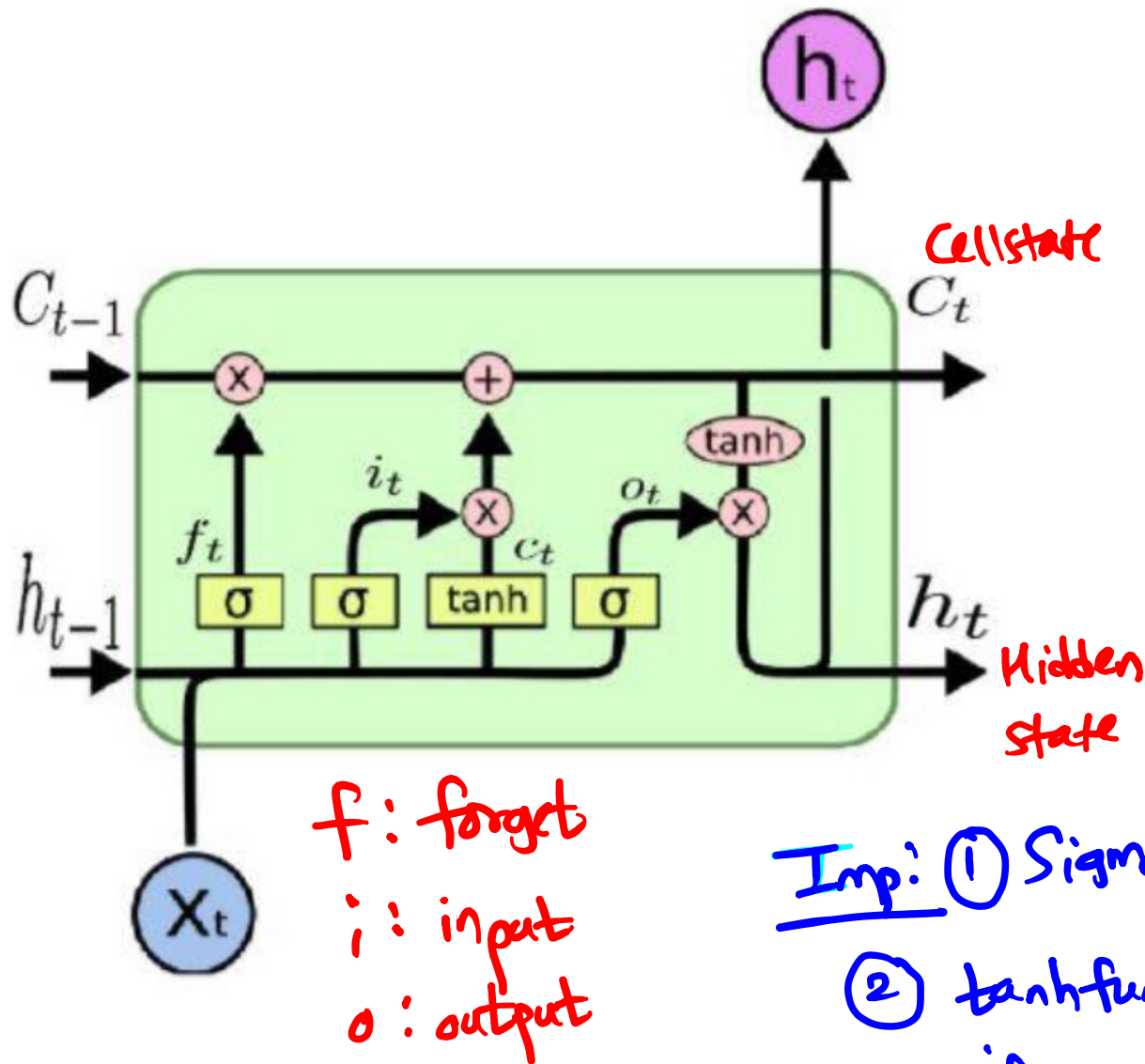
I live in Gujarat & I own a garment business here. I know ?. [Context - Far]

∴ We go for LSTM.

# LSTM Architecture



1. Forget gate: Whether to erase cell
2. Input gate: whether to write to cell
3. Gate gate: How much to write to cell
4. Output gate: How much to reveal cell.



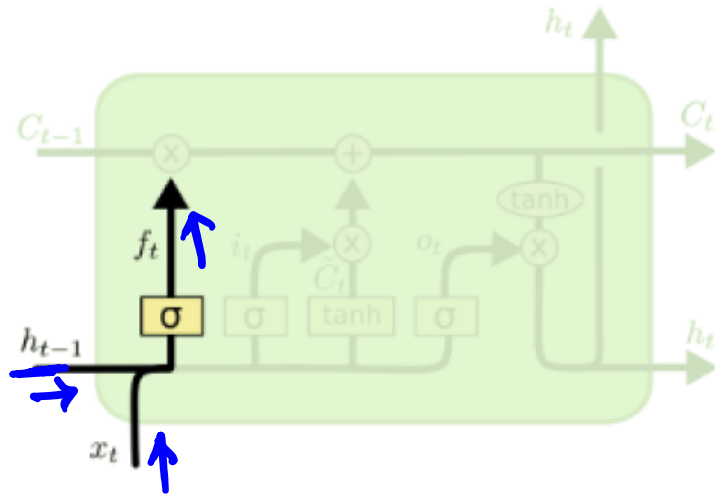
Two states that are transferred to next cell

1. Cellstate
2. Hidden State

Imp:

- ① Sigmoid function is always used to filter.
- ② tanh function is used for adding new or modifying existing information.

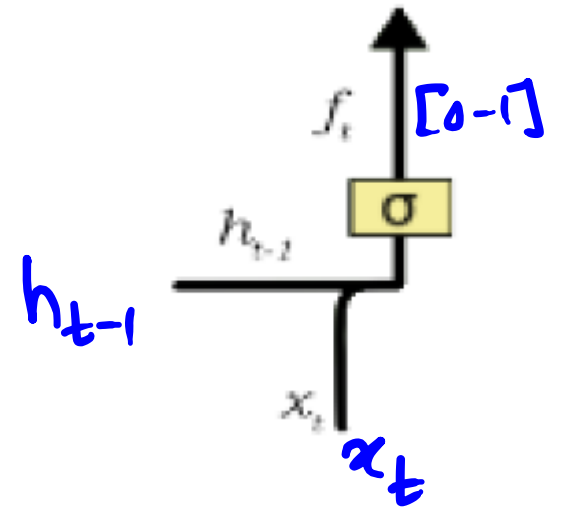
# Forget Gate



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

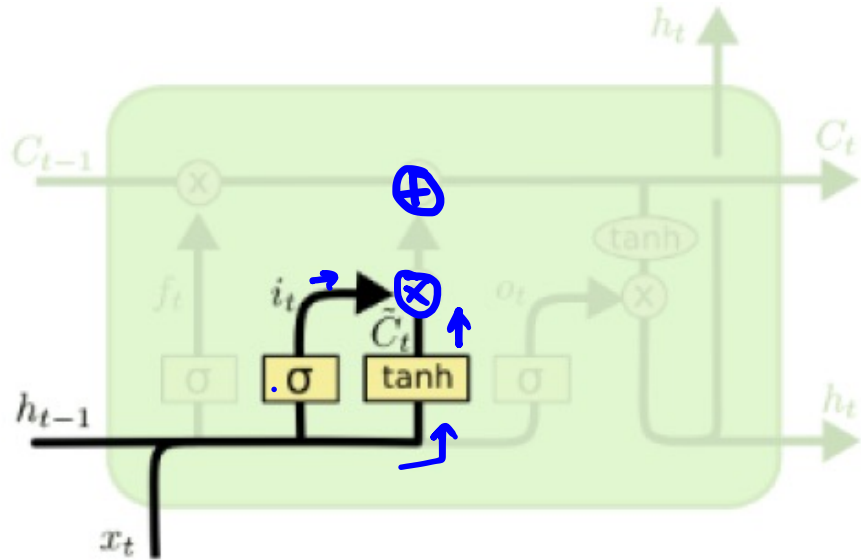
# Forget Gate Example

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



# Input Gate

we decide to add new stuff from present i/p to to  
our present cell state



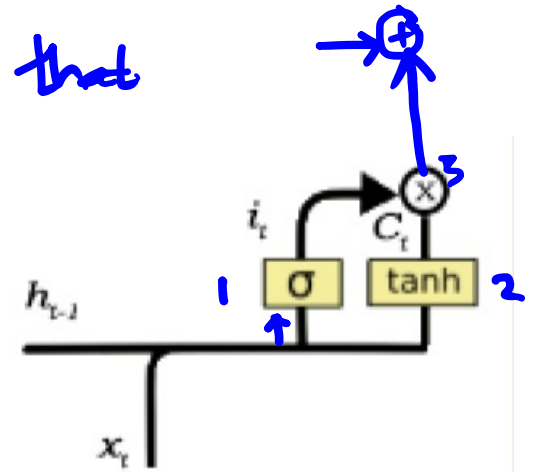
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

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

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

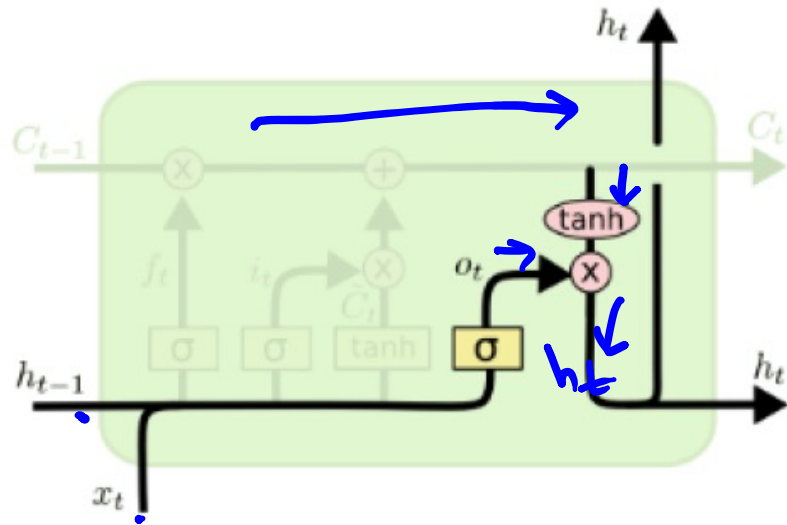
# Input Gate Example

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





Output Gate <sup>will</sup> We decide what to o/p from our cell state which will be done by our sigmoid func.

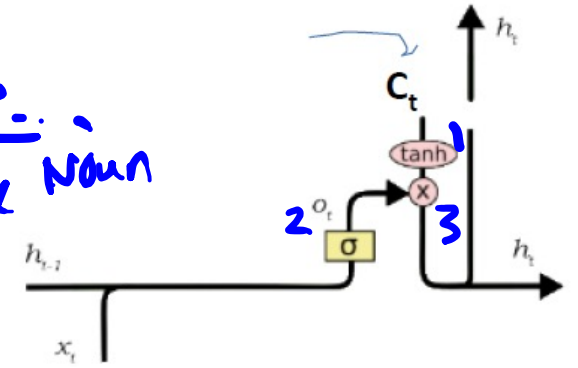


$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o)$$

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

# Output Gate Example

Bob fought single handedly with the enemy &  
died for his country. for his contributions brave Bob .  
↓  
Adjective Noun



# Additional Resource

- <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>