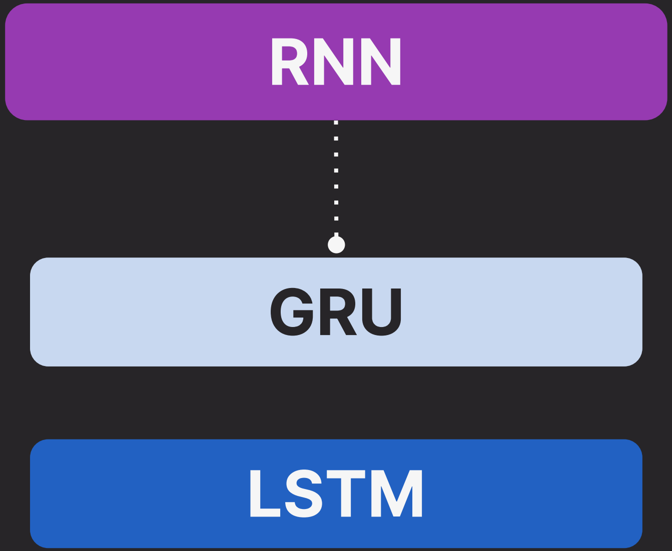
NLP: Recurrent Neural Network

Video 11: Understanding Long Short Term Memory Network

Analytics Vidhya

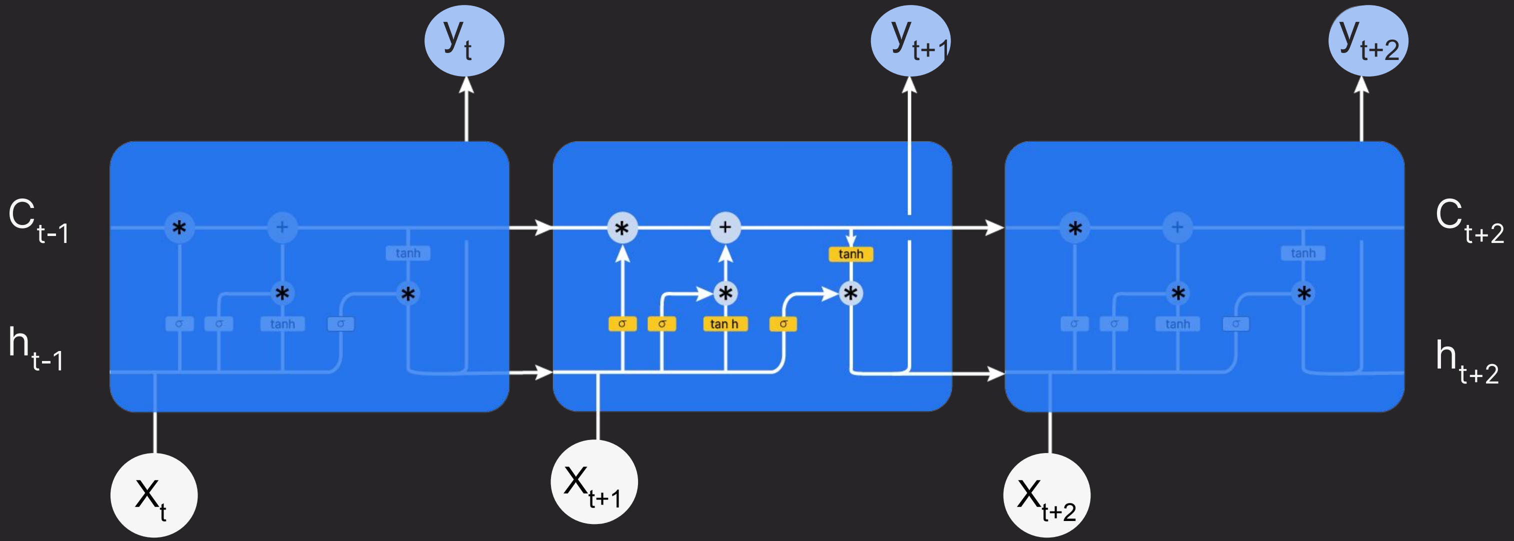
Advanced RNN



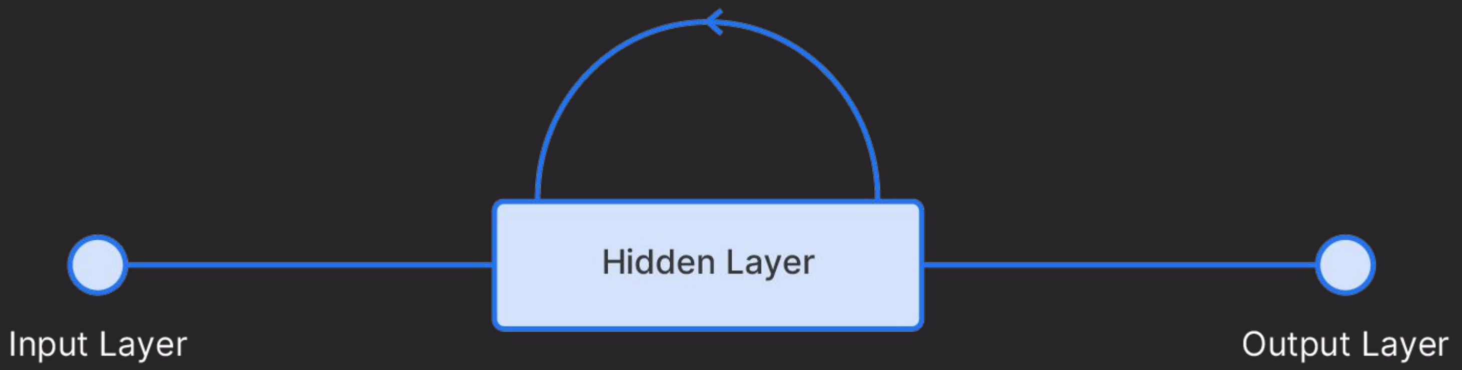
Introduction to LSTM

Analytics Vidhya

Long Short Term Memory (LSTM) excels at capturing long-term dependencies making it ideal for sequence prediction.

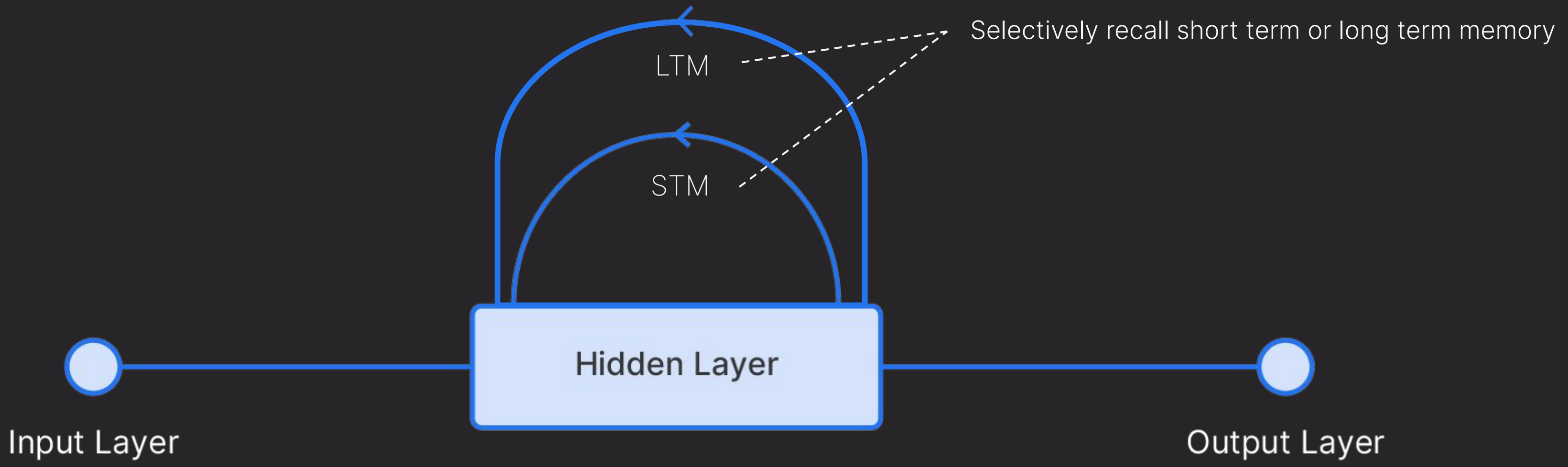


Introduction to LSTM



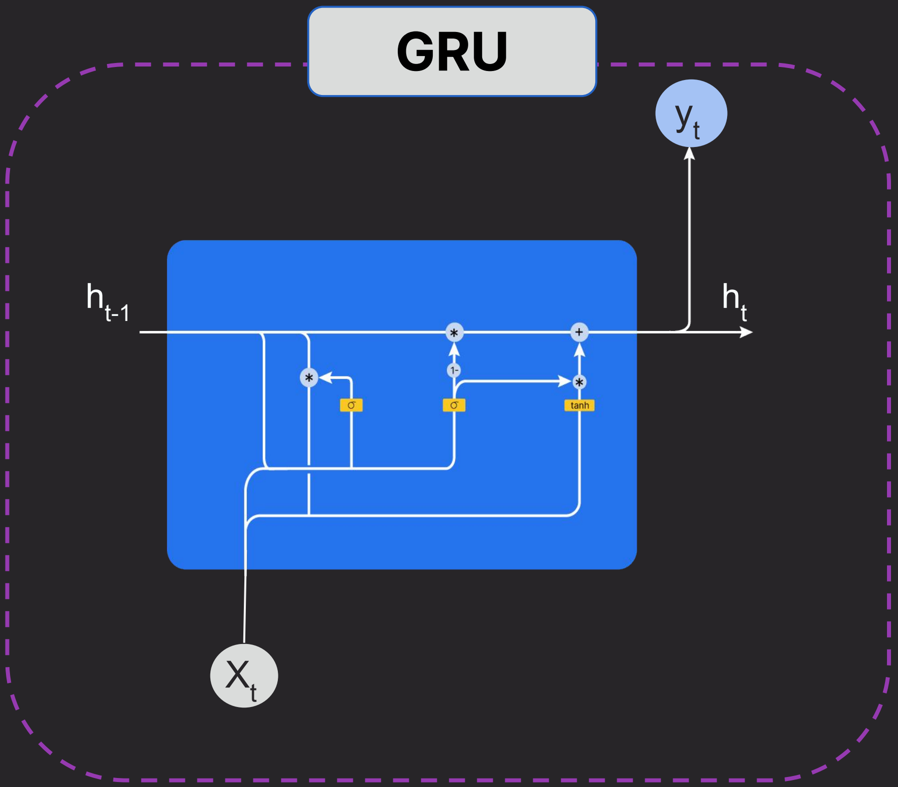
Structure of a Simple RNN

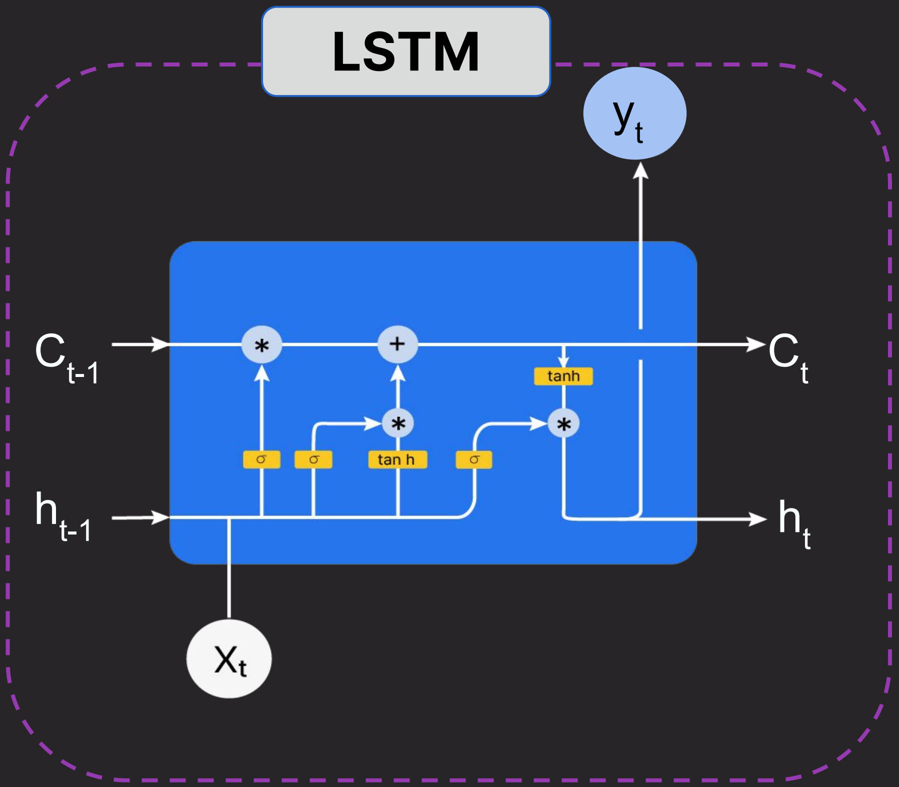
Introduction to LSTM



Structure of LSTM

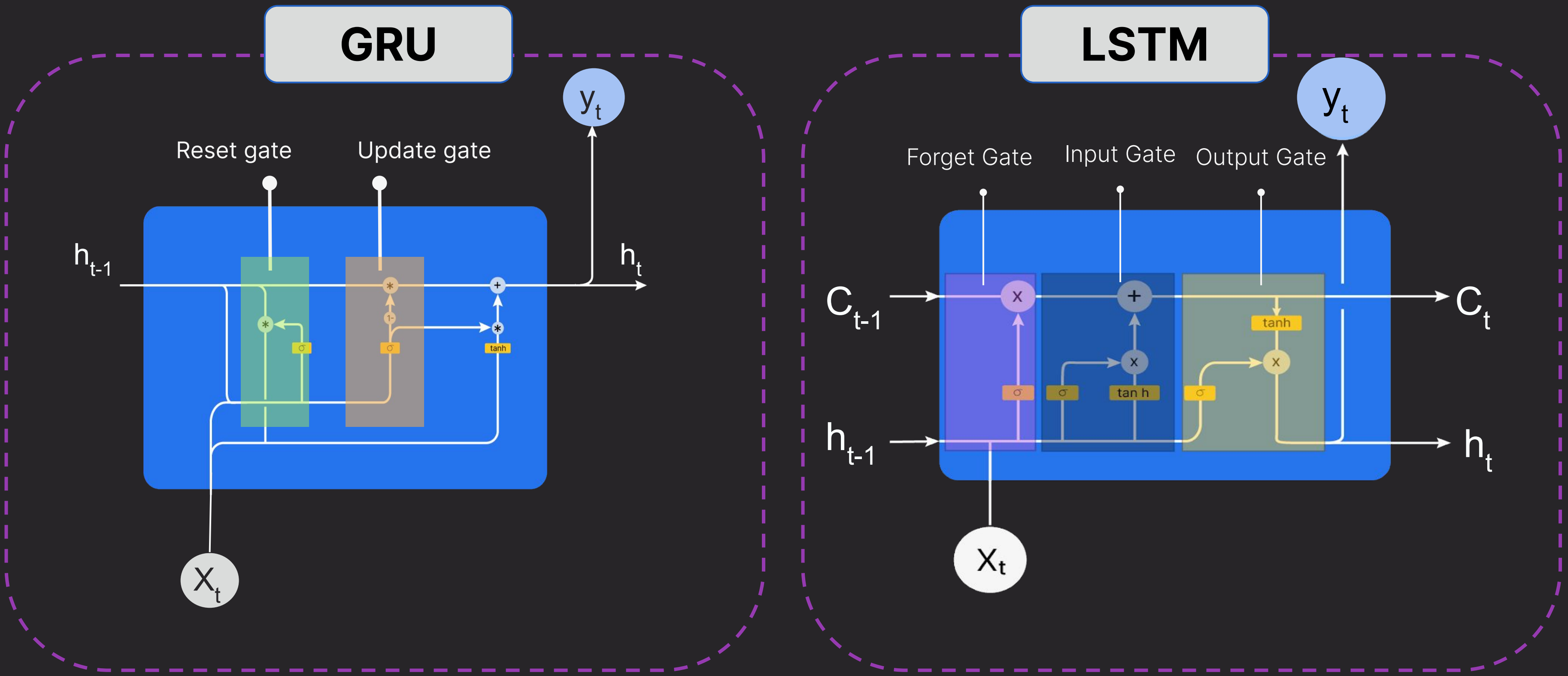
RNN V/s GRU





RNN V/s GRU

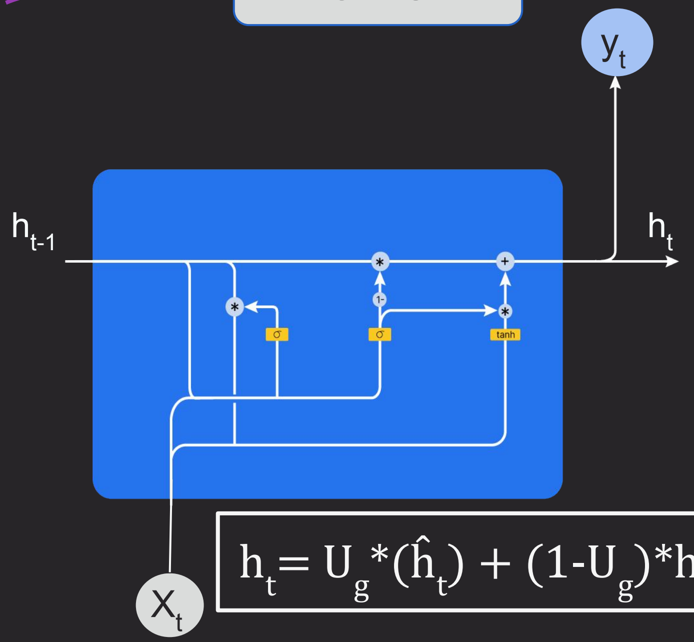
Analytics Vidhya



Analytics Vidhya

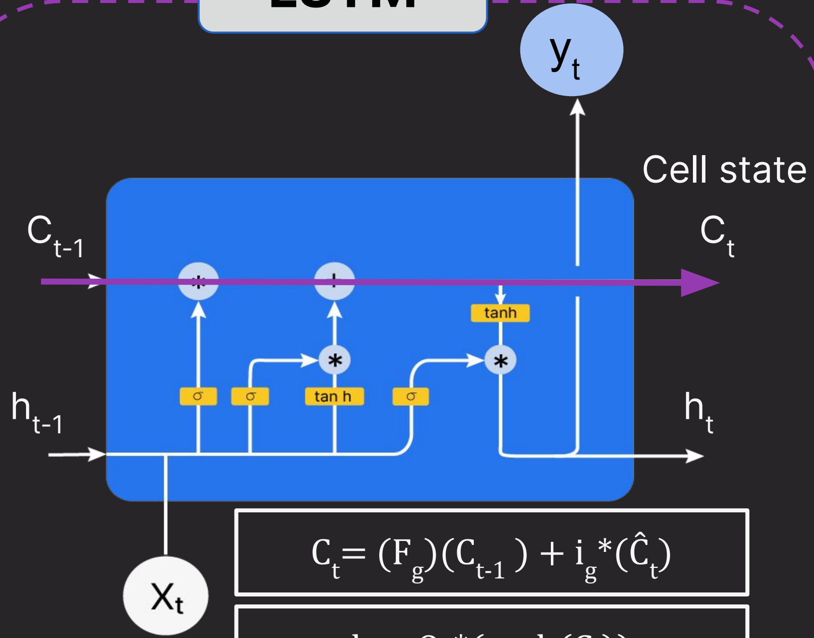
RNN V/s GRU

GRU



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

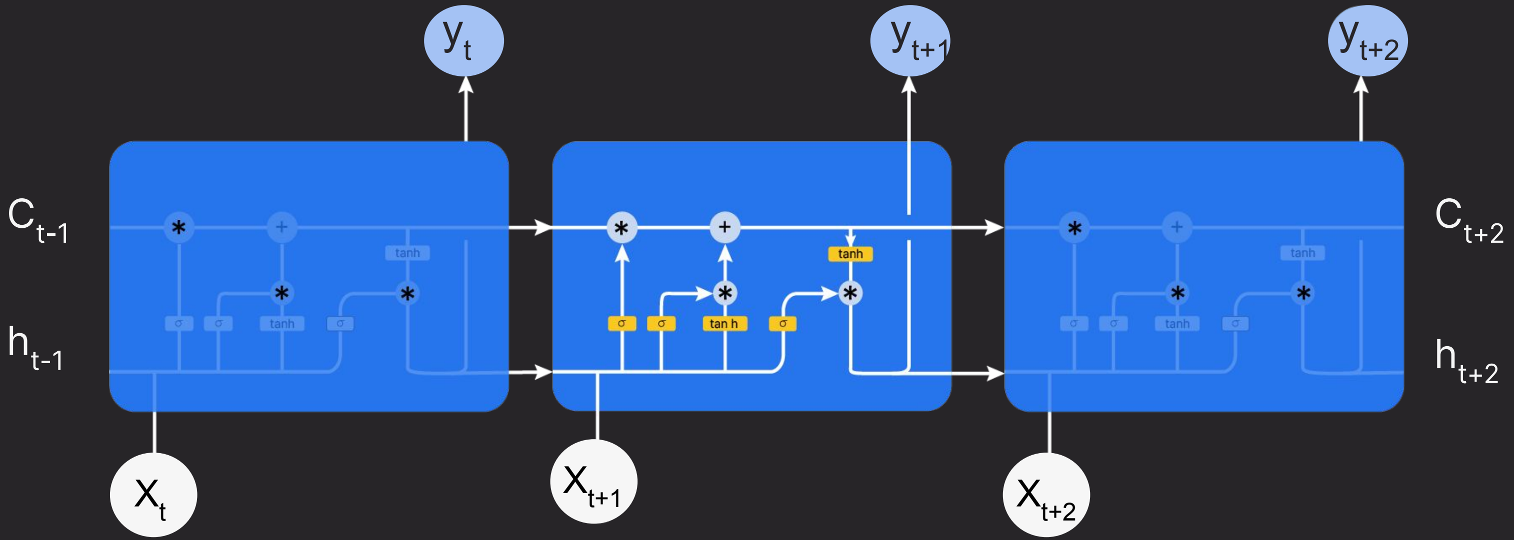
LSTM



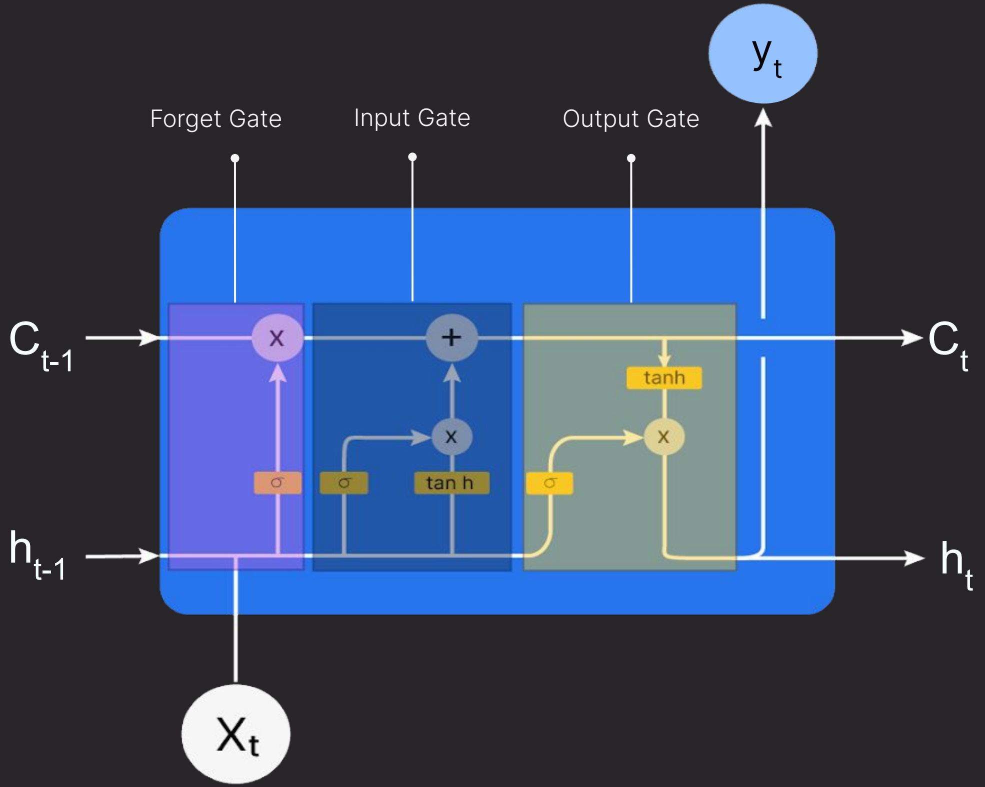
$$C_t = (F_g)(C_{t-1}) + i_g * (\hat{C}_t)$$

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

LSTM Architecture

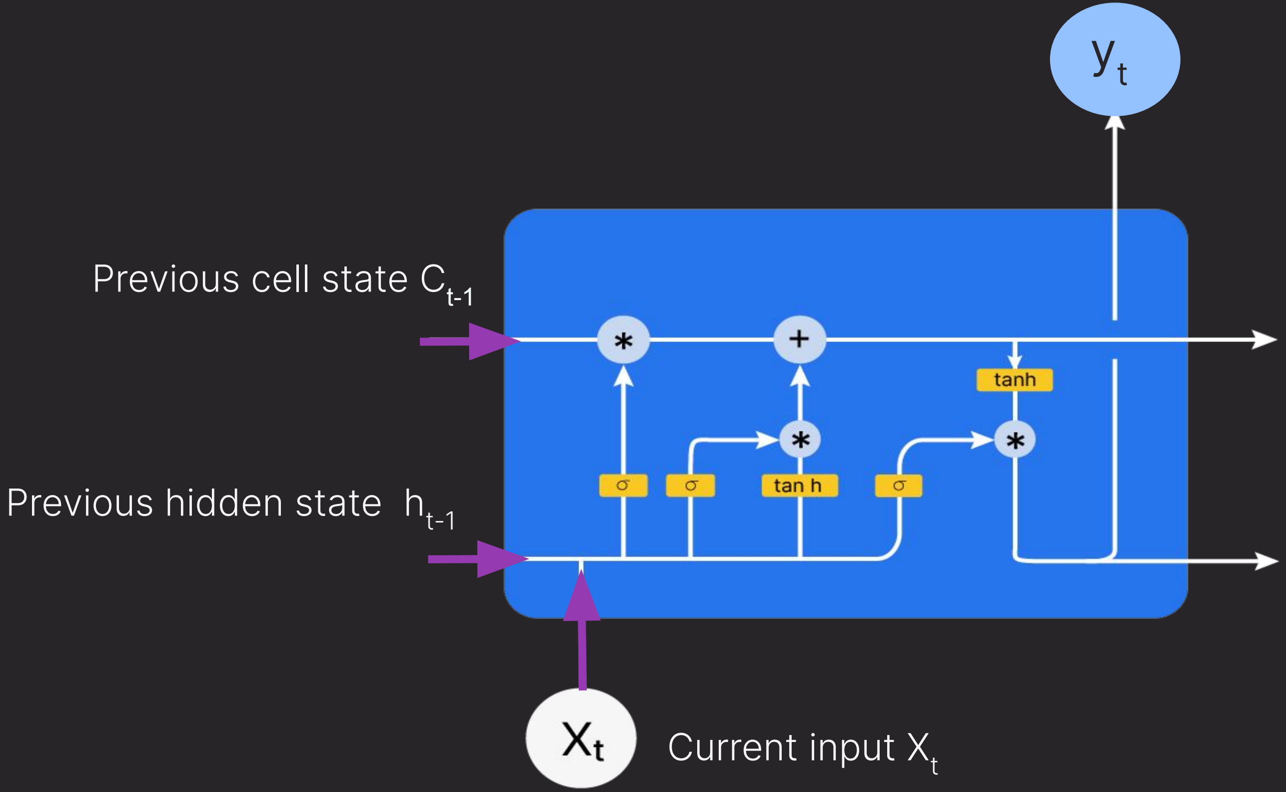


LSTM Architecture

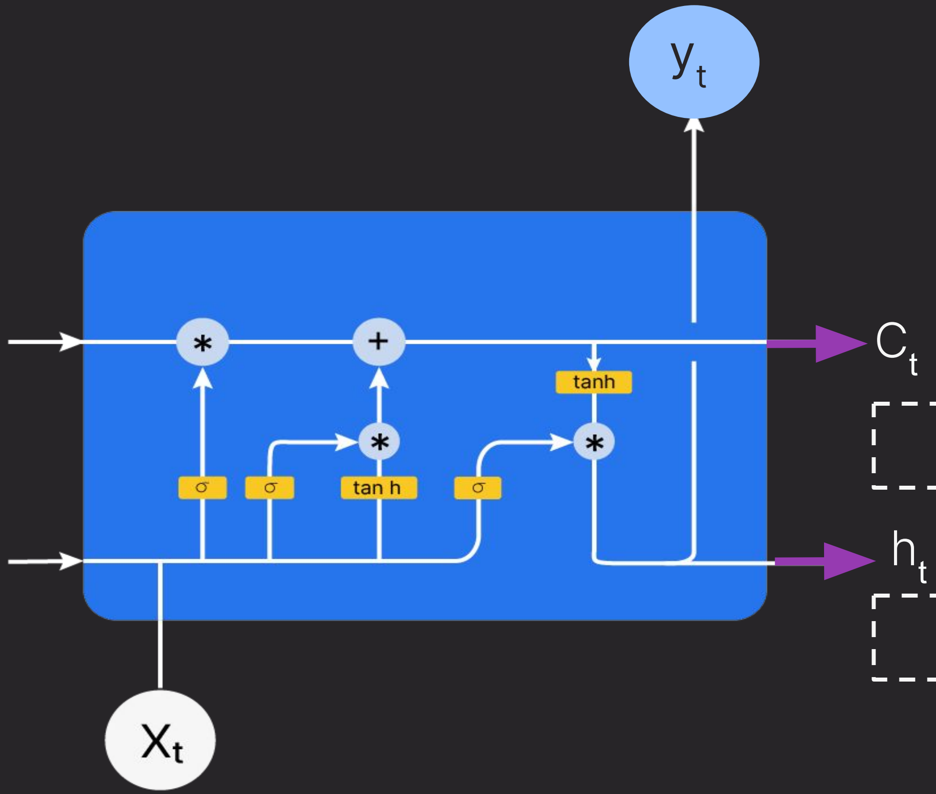


- **Forget Gate:** Decides informations to discard from the cell state
- **Input Gate:** Decides new information to add to the cell state.
- **Output Gate:** Determines the current time step's hidden state (h_t).

Understanding LSTM Process



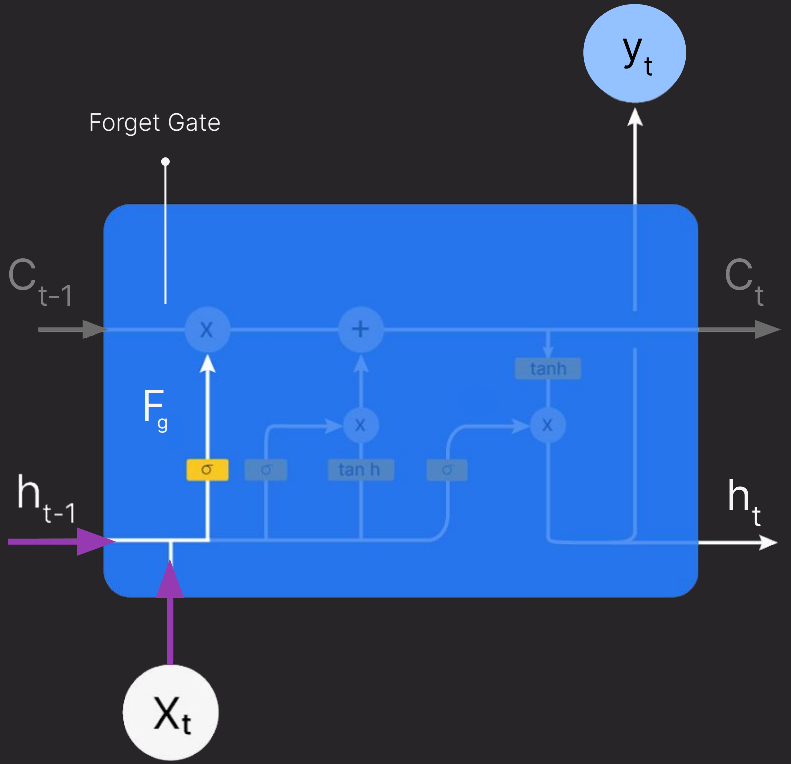
Understanding LSTM Process



$$C_t = (F_g)(C_{t-1}) + i_g * (\hat{C}_t)$$

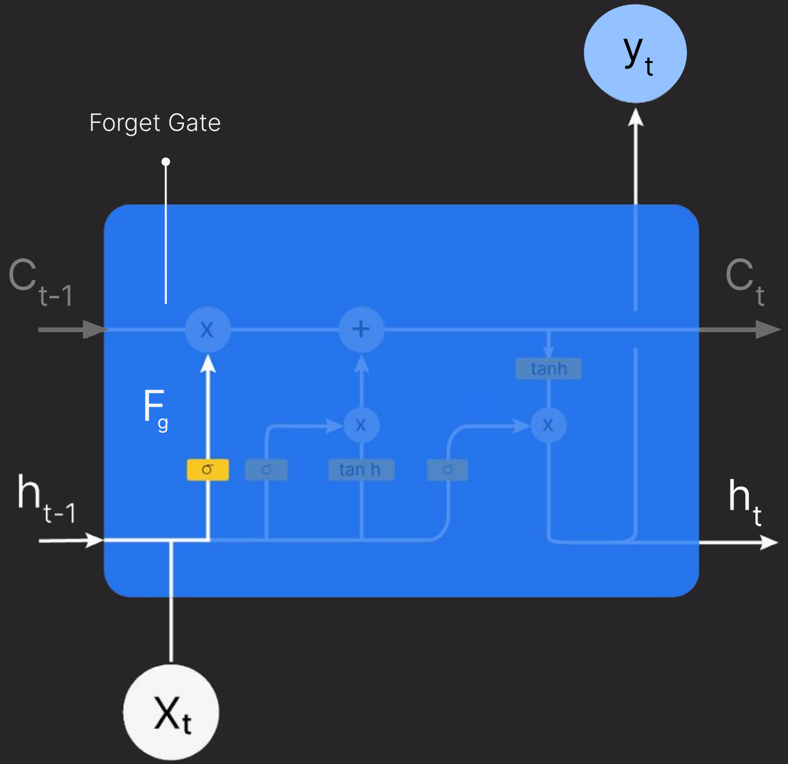
$$h_t = O_g * (\tanh(C_t))$$

Understanding LSTM Process



$$F_g = \sigma(W_f [X_t , h_{t-1}]+ b_f)$$

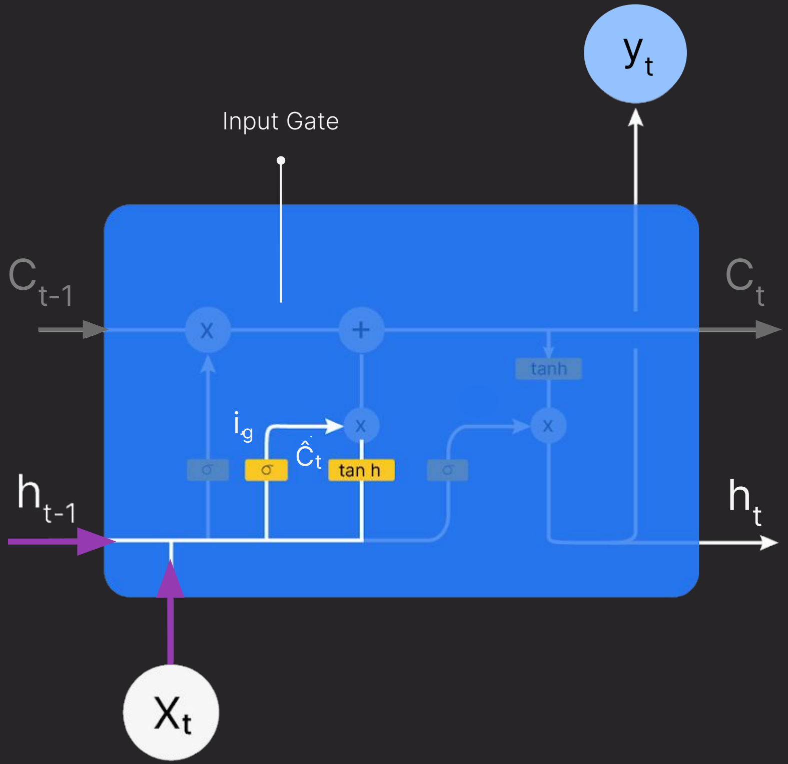
Understanding LSTM Process



$$F_g = \sigma(W_f \, [X_t , h_{t-1}] + b_f)$$

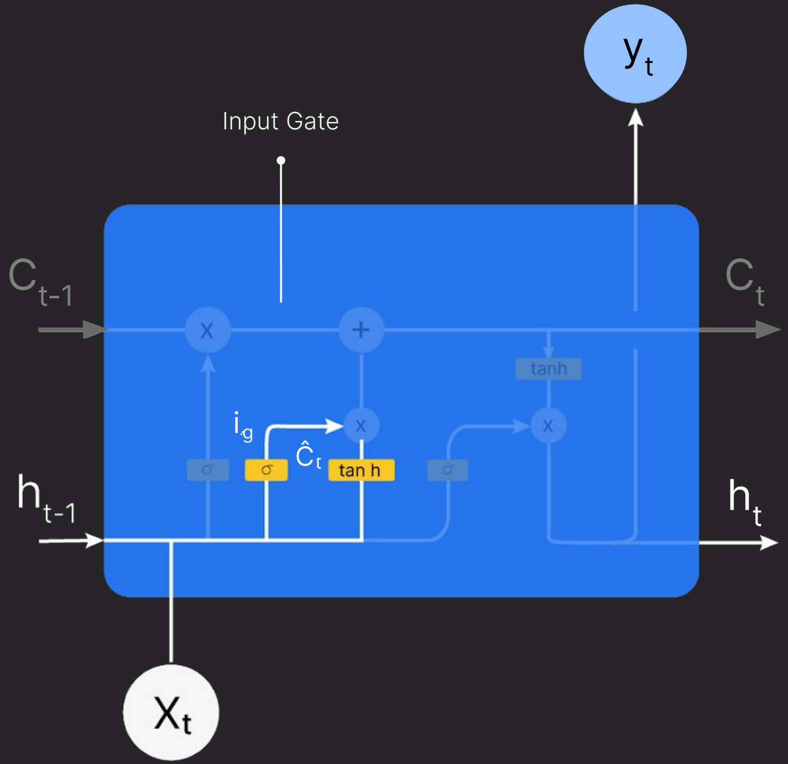
- **Outputs** → Between 0 and 1
- **0** → “Forget this completely”
- **1** → “Keep this entirely”

Understanding LSTM Process



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

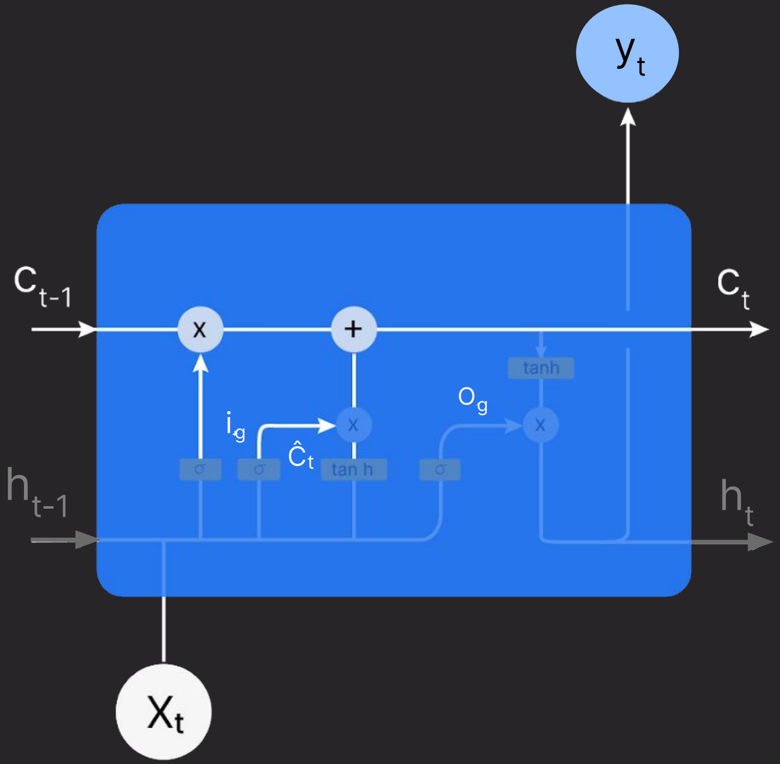
Understanding LSTM Process



New candidate Vector

$$\hat{c}_t = \tanh(W_c [x_t + h_{t-1}] + b_c)$$

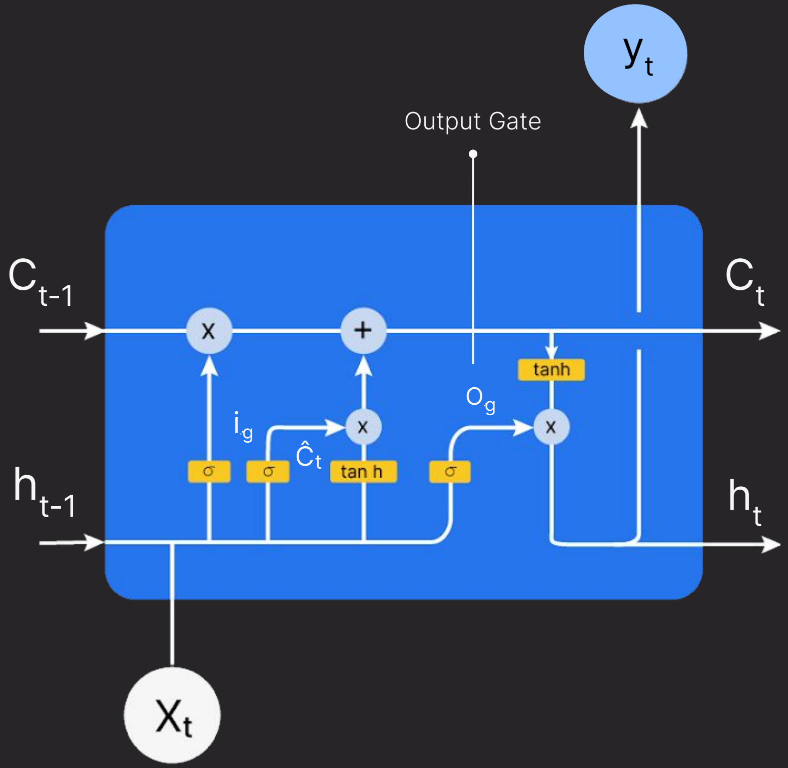
Understanding LSTM Process



Update the cell state :

$$C_t = (F_g)(C_{t-1}) + i_g * (\hat{c}_t)$$

Understanding LSTM Process



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

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

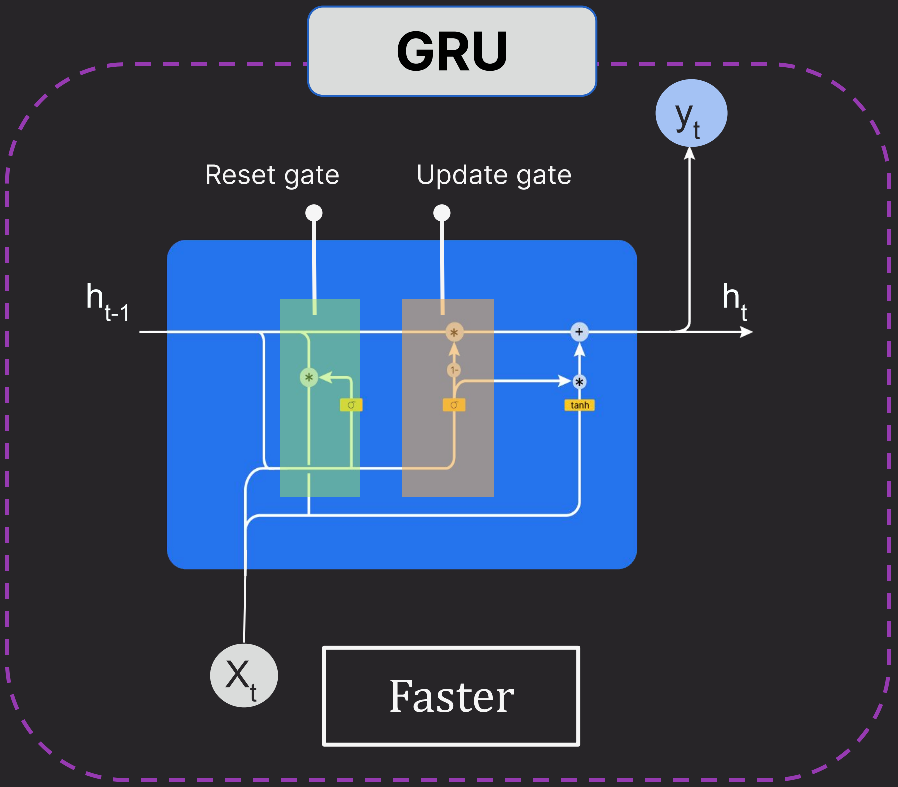
Advanced RNN

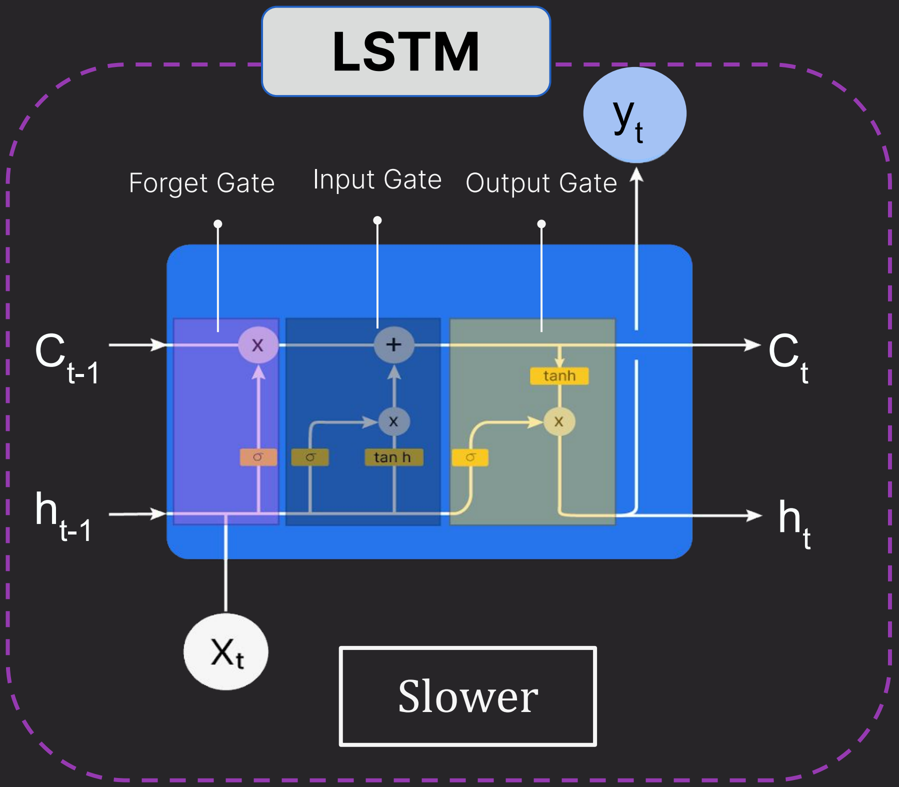
RNN

GRU

LSTM

RNN V/s GRU


GRU
Reset gate
Update gate
h_{t-1}
h_t
y_t
tanh
X_t
Faster


LSTM
Forget Gate
Input Gate
Output Gate
y_t
C_{t-1}
C_t
h_{t-1}
h_t
tanh
tan h
X_t
Slower

Up Next: LSTM in Jupyter