NLP: Recurrent Neural Network

Video 9: Understanding GRUs

Analytics Vidhya

Analytics Vidhya

Advanced RNN

RNN

GRU

LSTM

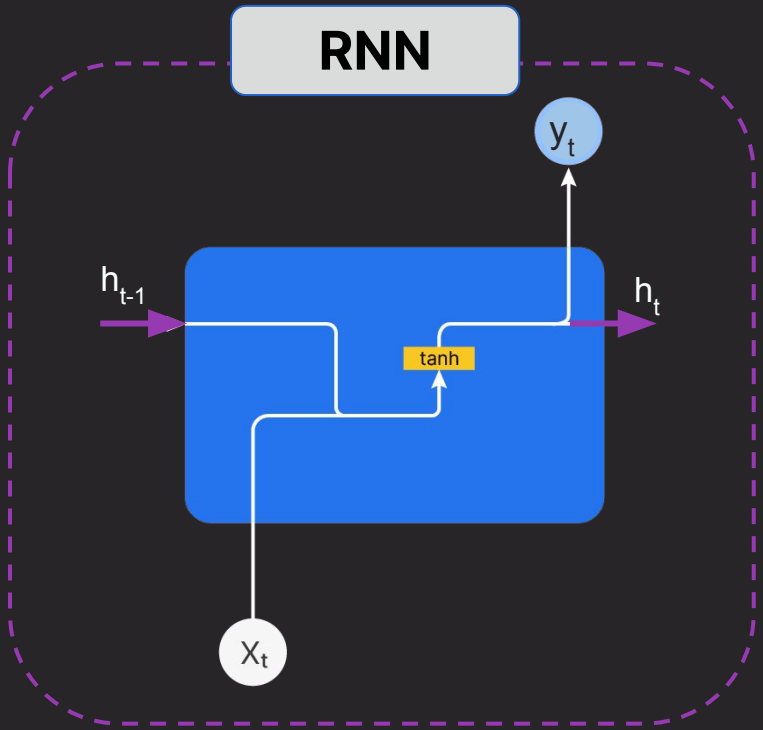
Advanced RNN

RNN

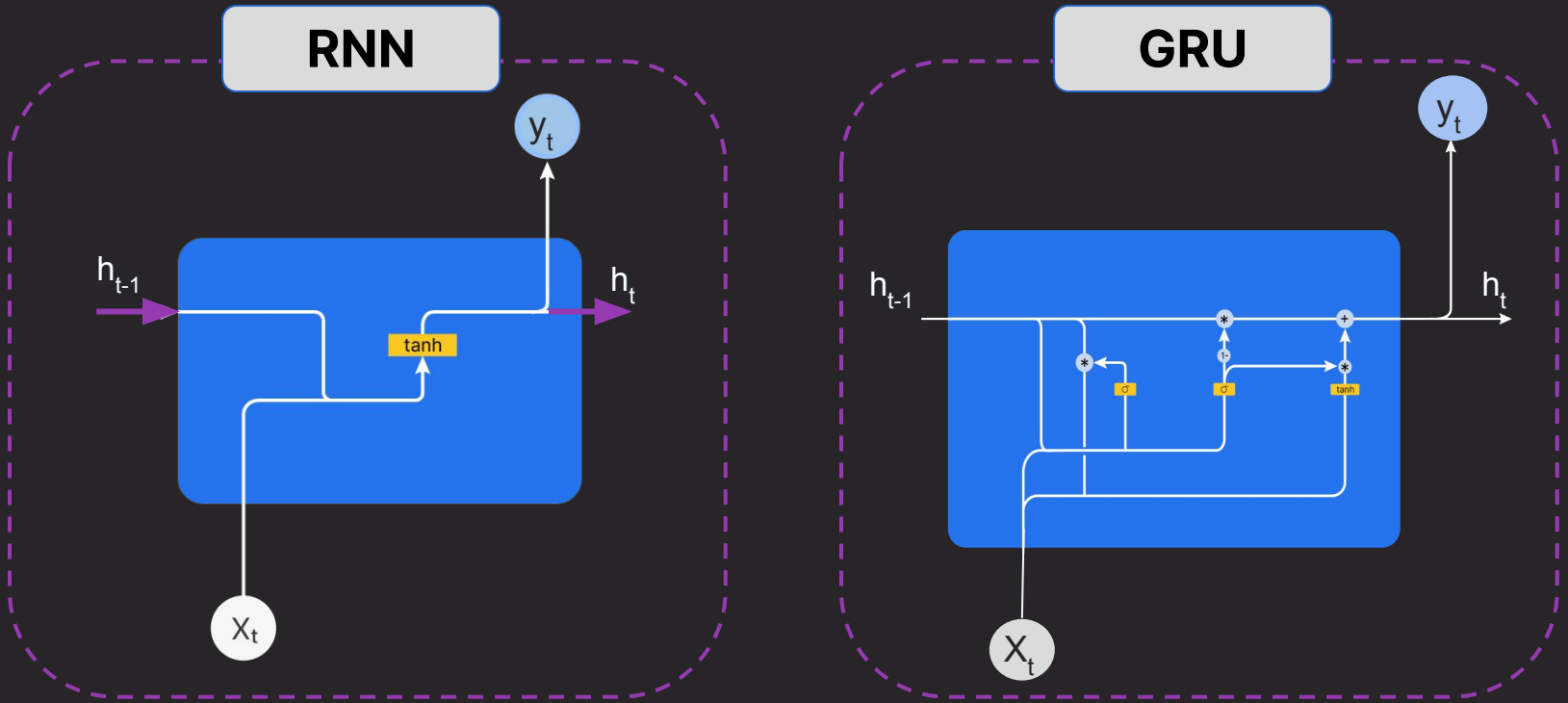
GRU

LSTM

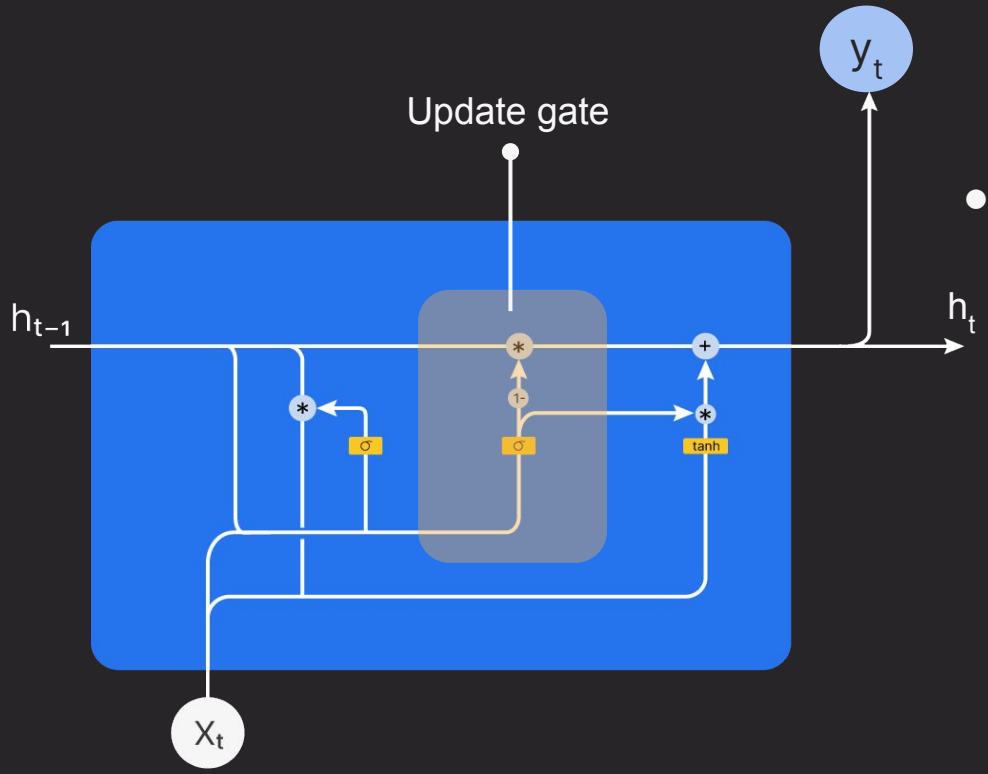
RNN



RNN V/s GRU



Architecture of GRU



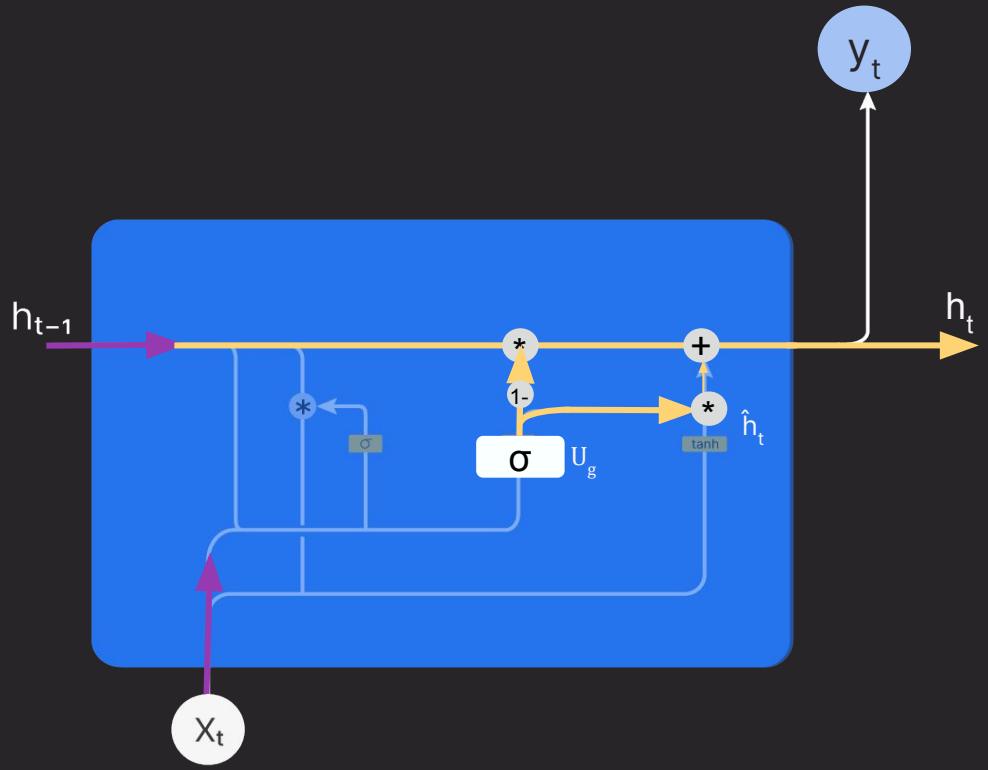
- Update gate : Select component from previous and current hidden state

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

Candidate hidden state

Update gate

GRU : Update gate

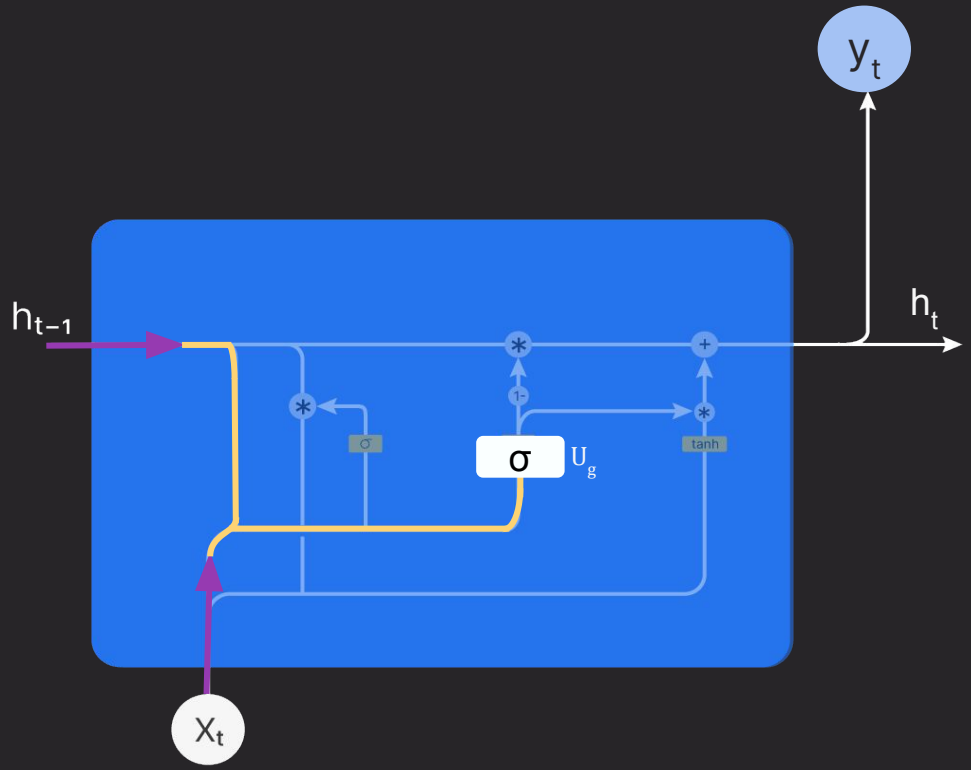


$$h_t = U_g * (\hat{h}_t) + (1\text{-}U_g) * h_{t-1}$$

U_g ····· 1 Prioritize new information

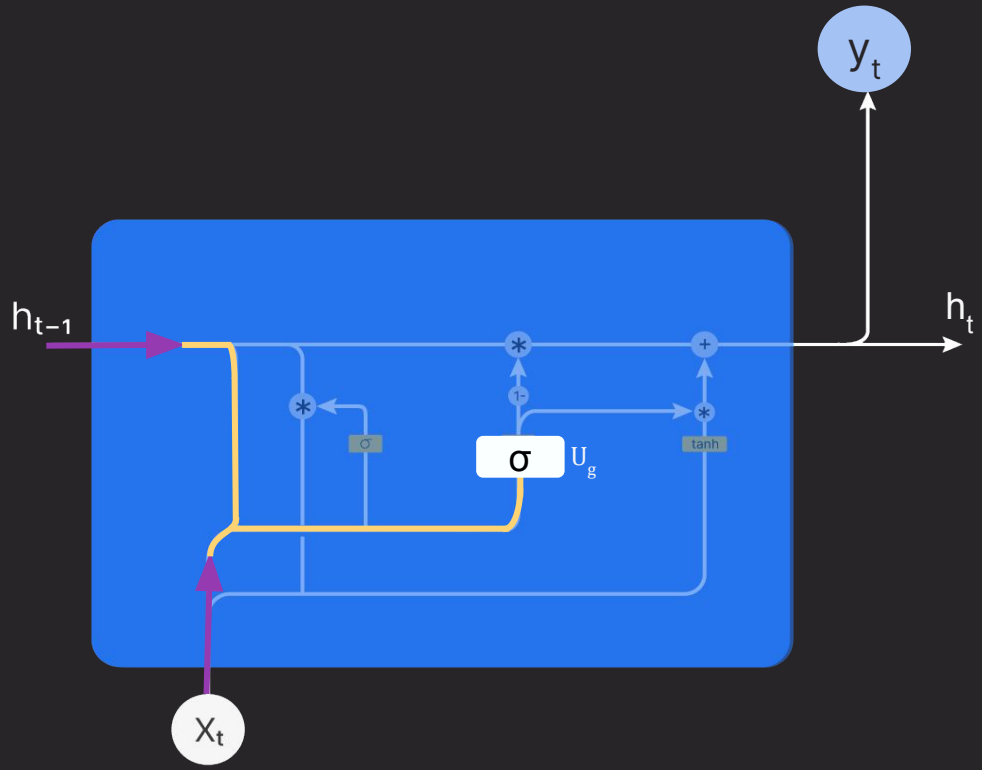
U_g ····· 0 Retain old information

GRU : Calculation of Update gate



$$U_g = \sigma(X_t . W_{xu} + h_{t-1} . W_{hu} + b_u)$$

GRU : Calculation of Update gate

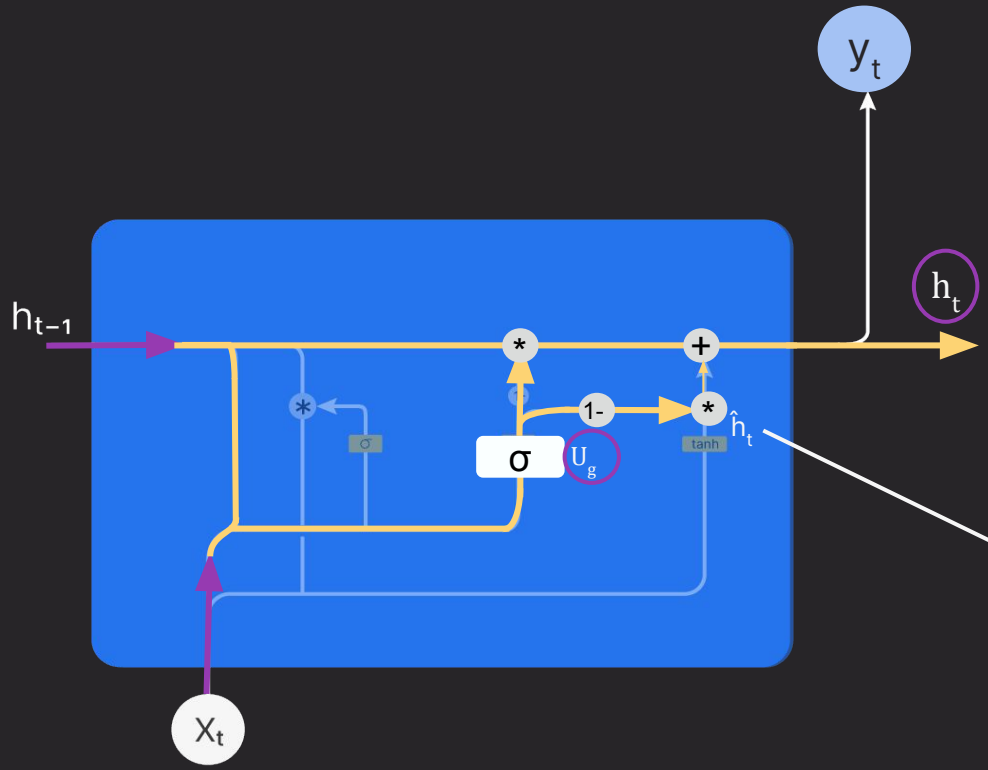


W_u

$$U_g = \sigma(X_t . W_{xu} + h_{t-1} . W_{hu} + b_u)$$

$$U_g = \sigma(W_u [X_t , h_{t-1}] + b_u)$$

GRU : Update gate

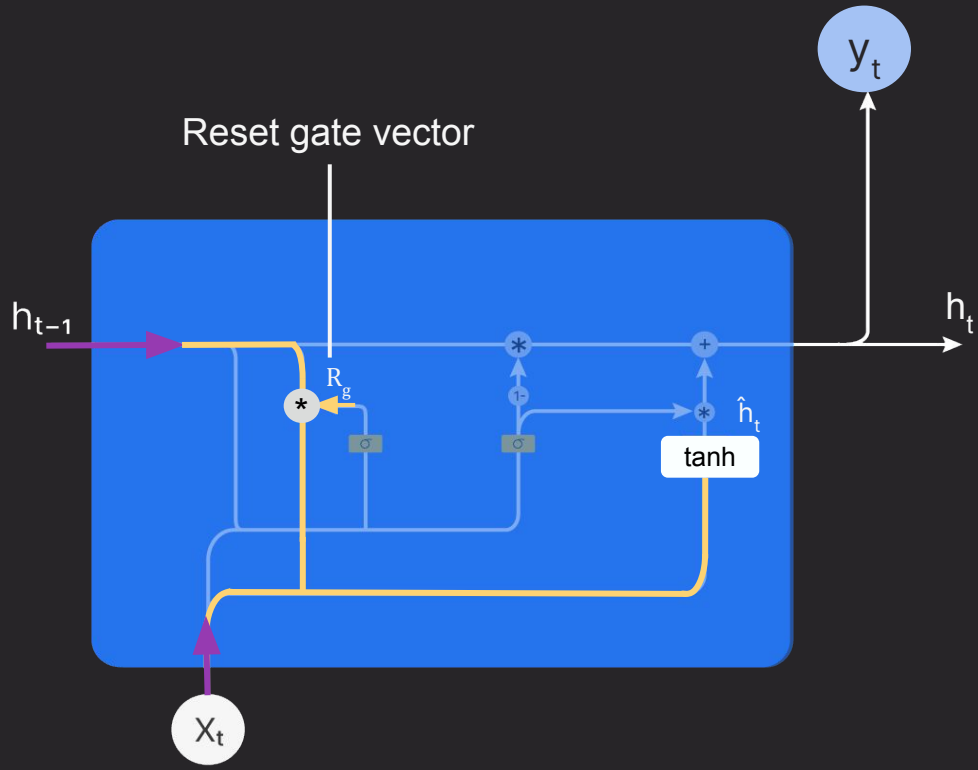


$$U_g = \sigma(W_u [X_t , h_{t-1}] + b_u)$$

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

Candidate hidden state

GRU : Calculation of Candidate hidden state

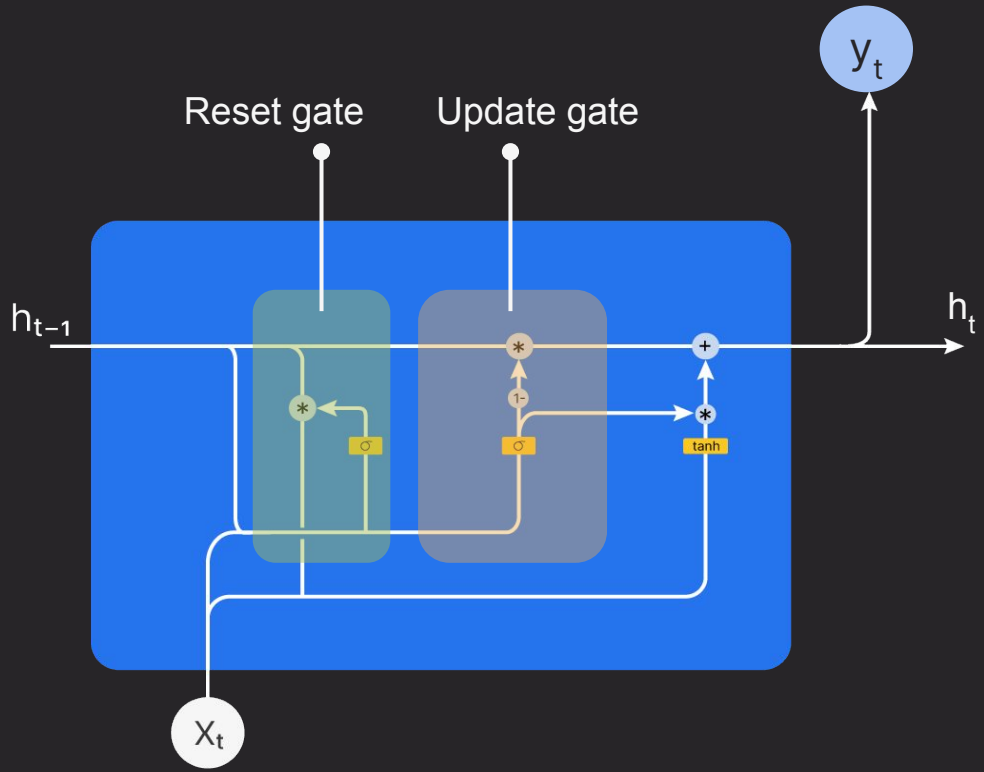


- Candidate hidden state : Represent the potential new content to be part of current hidden state.

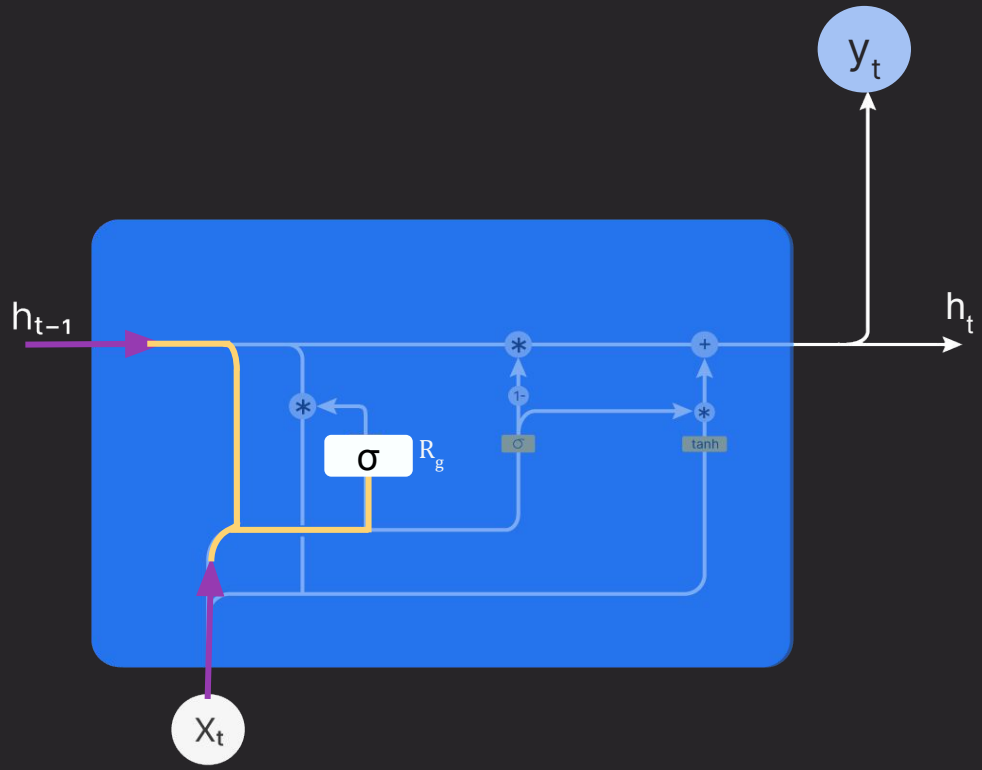
$$\hat{h}_t = \tanh\left(W_h\left[x_t, R_g * h_{t-1}\right] + b_h\right)$$

Reset gate

Architecture of GRU

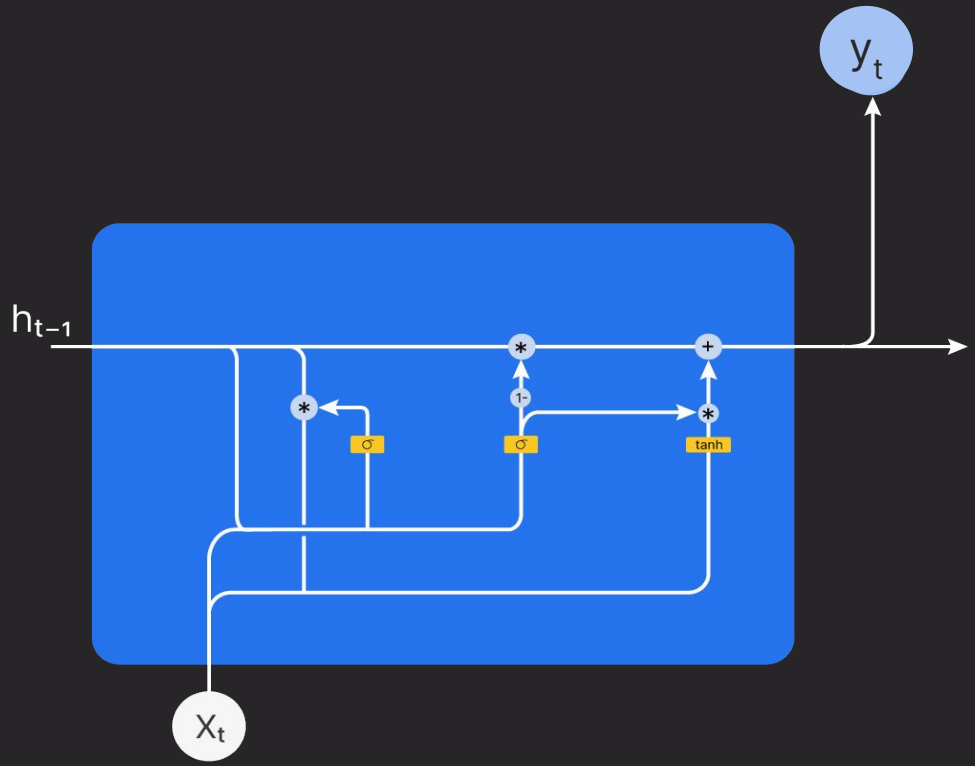


GRU : Calculation of Candidate hidden state



$$R_g = \sigma(W_R [X_t , h_{t-1}] + b_R)$$

Architecture of GRU



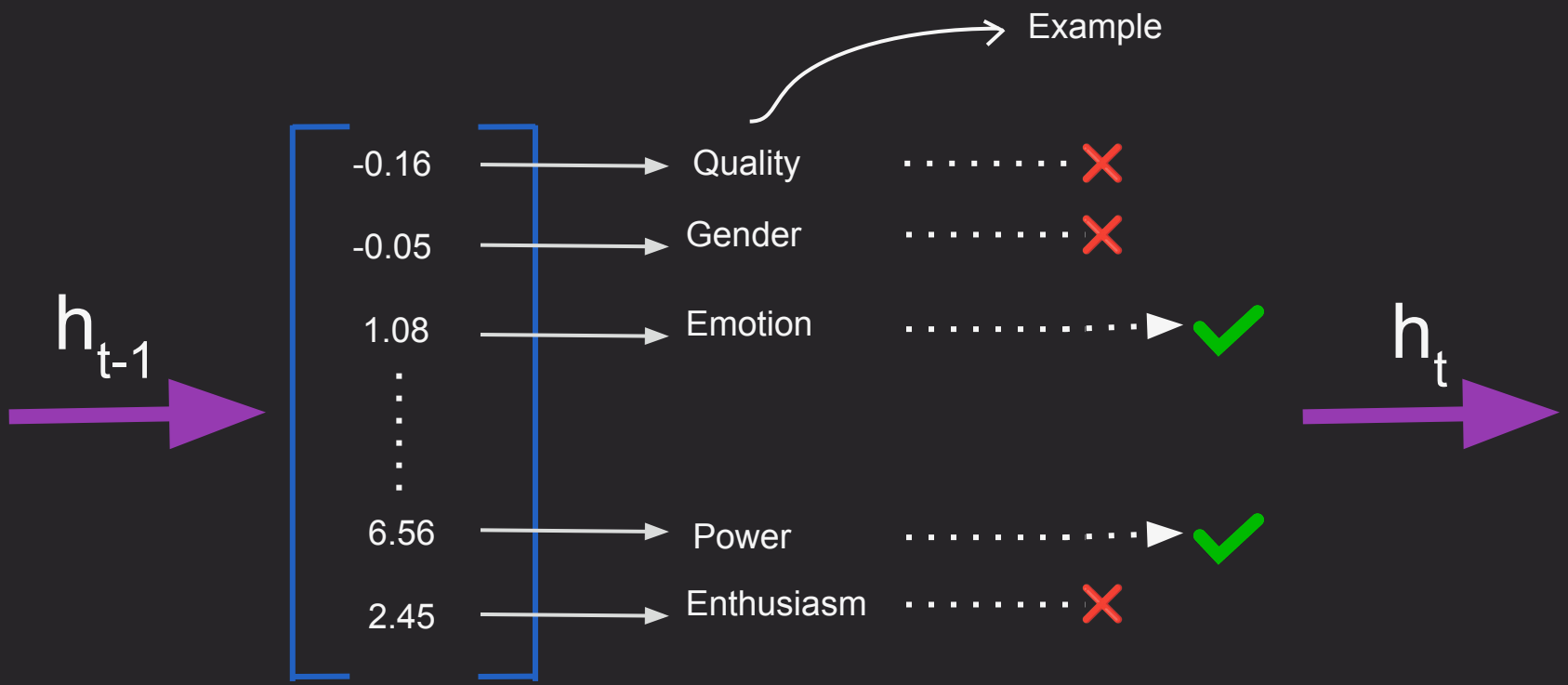
$$U_g = \sigma(W_u [X_t , h_{t-1}] + b_u)$$

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

$$\hat{h}_t = \tanh (W_h [x_t , R_g * h_{t-1}] + b_h)$$

$$R_g = \sigma(W_R [X_t , h_{t-1}] + b_R)$$

Architecture of GRU



Analytics Vidhya

Up Next: GRU in Jupyter