

Deep Learning

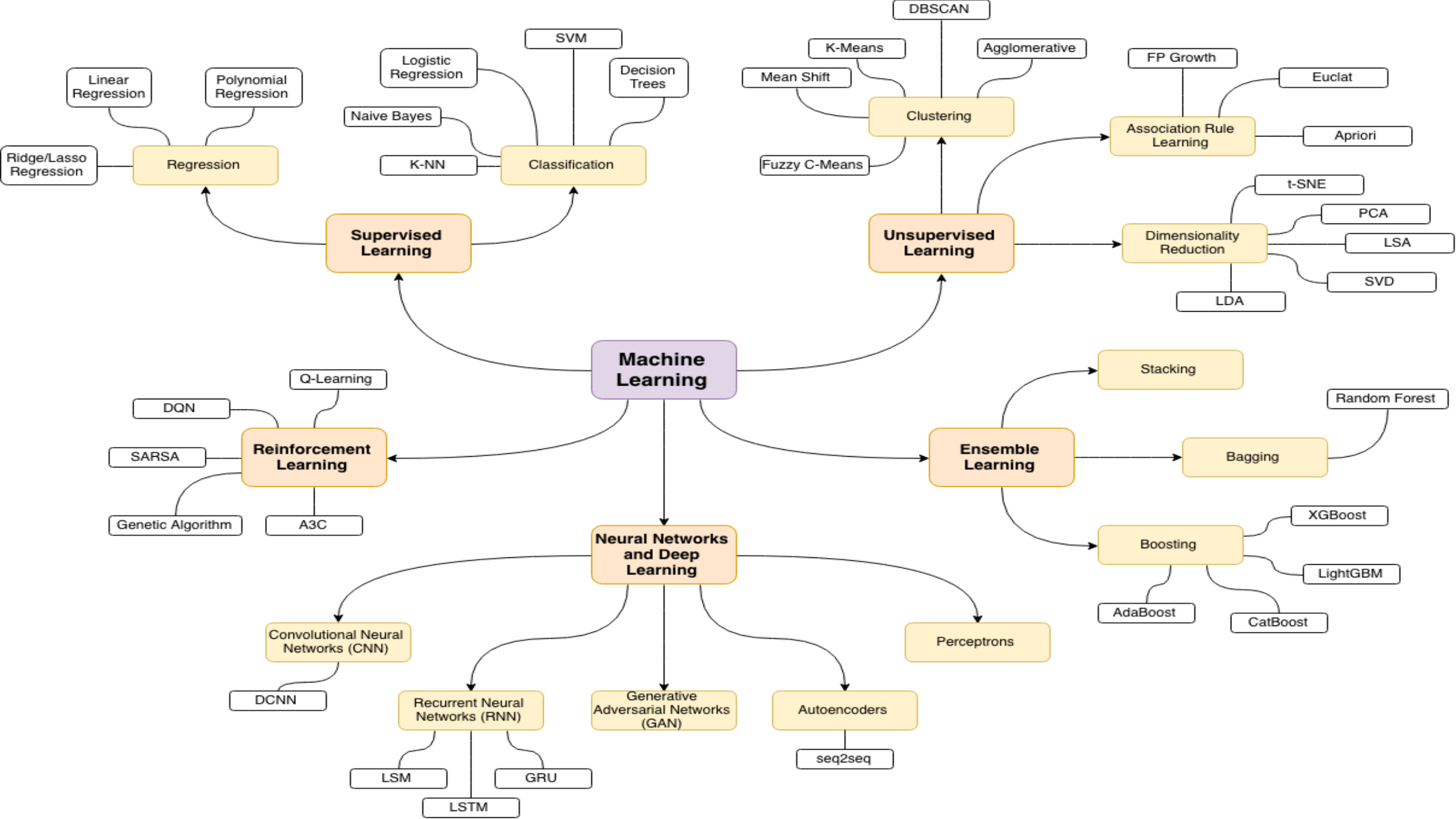
By MXK

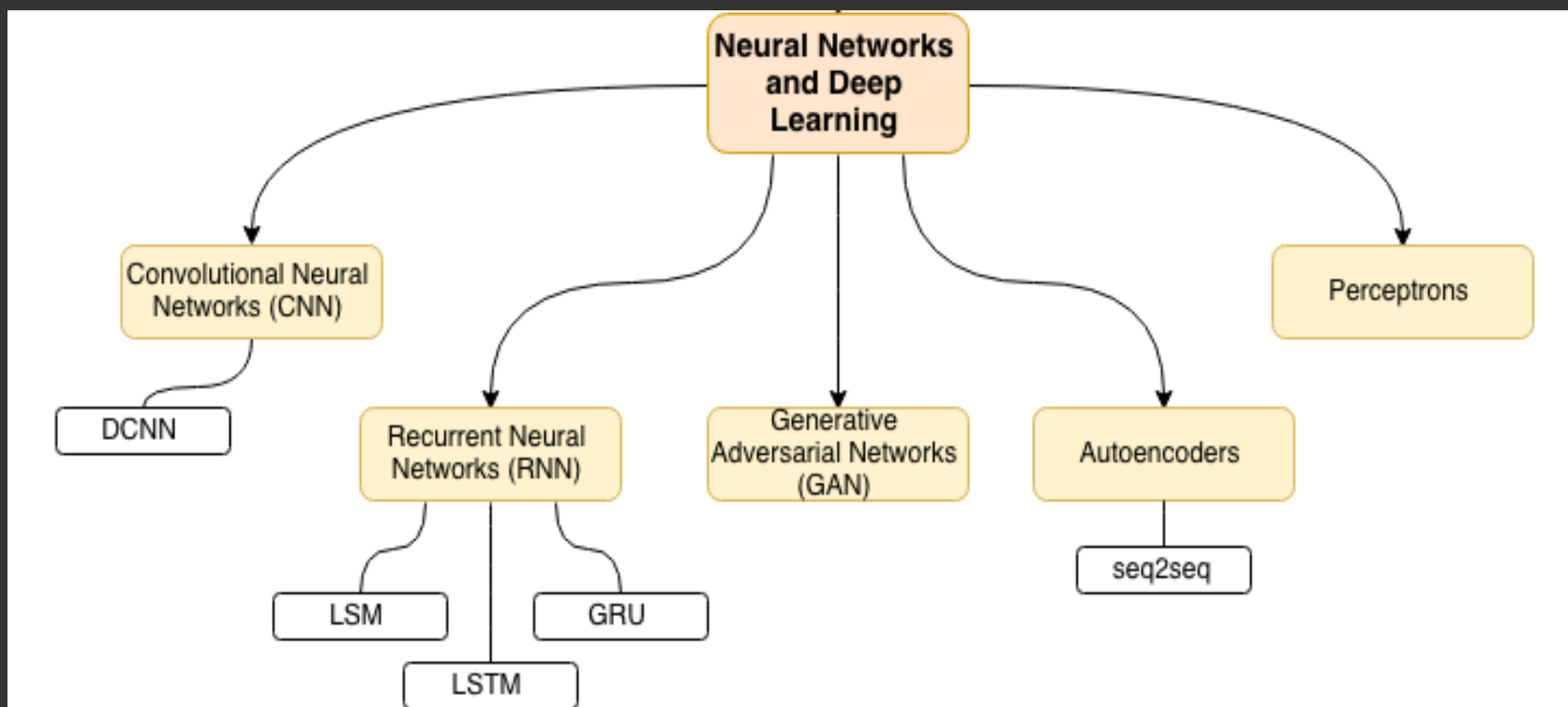


Summary

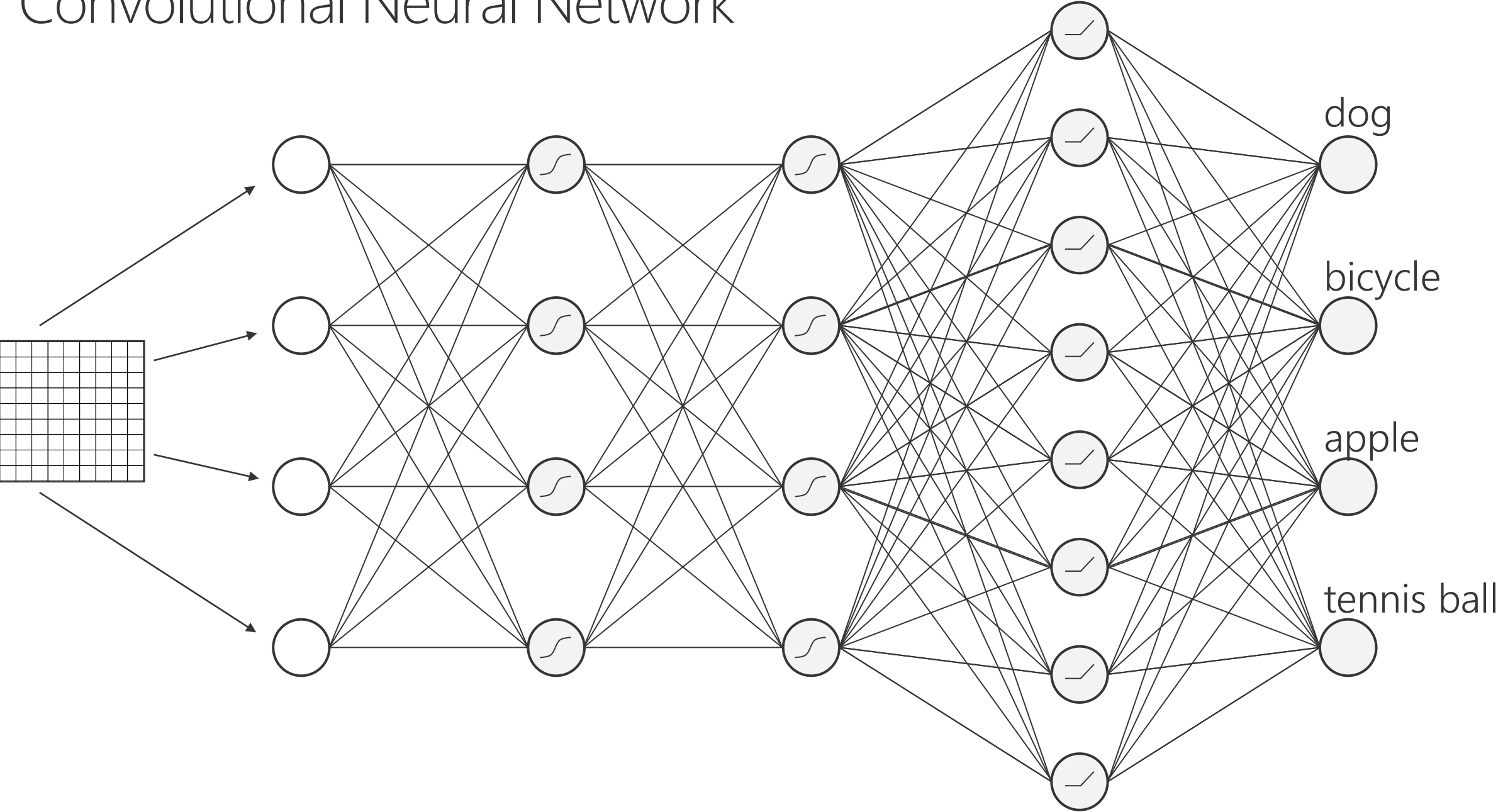
- 1- Machine Learning Roadmap
- 2- Deep Learning Roadmap
- 3- Convolutional Neural Network
- 4- Recurrent Neural Network
- 5- Long Short Term Memory
- 6- Closing Thoughts



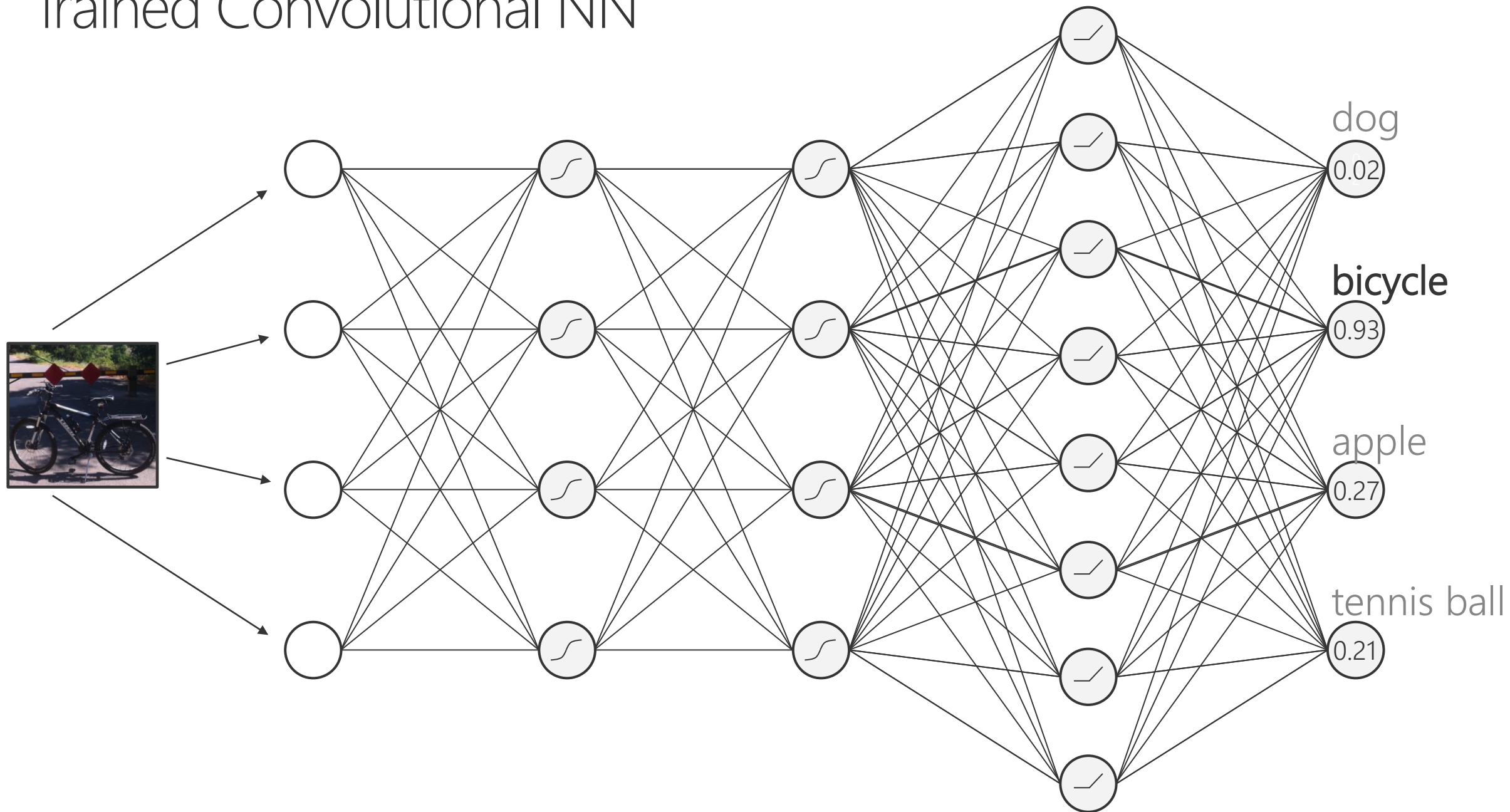




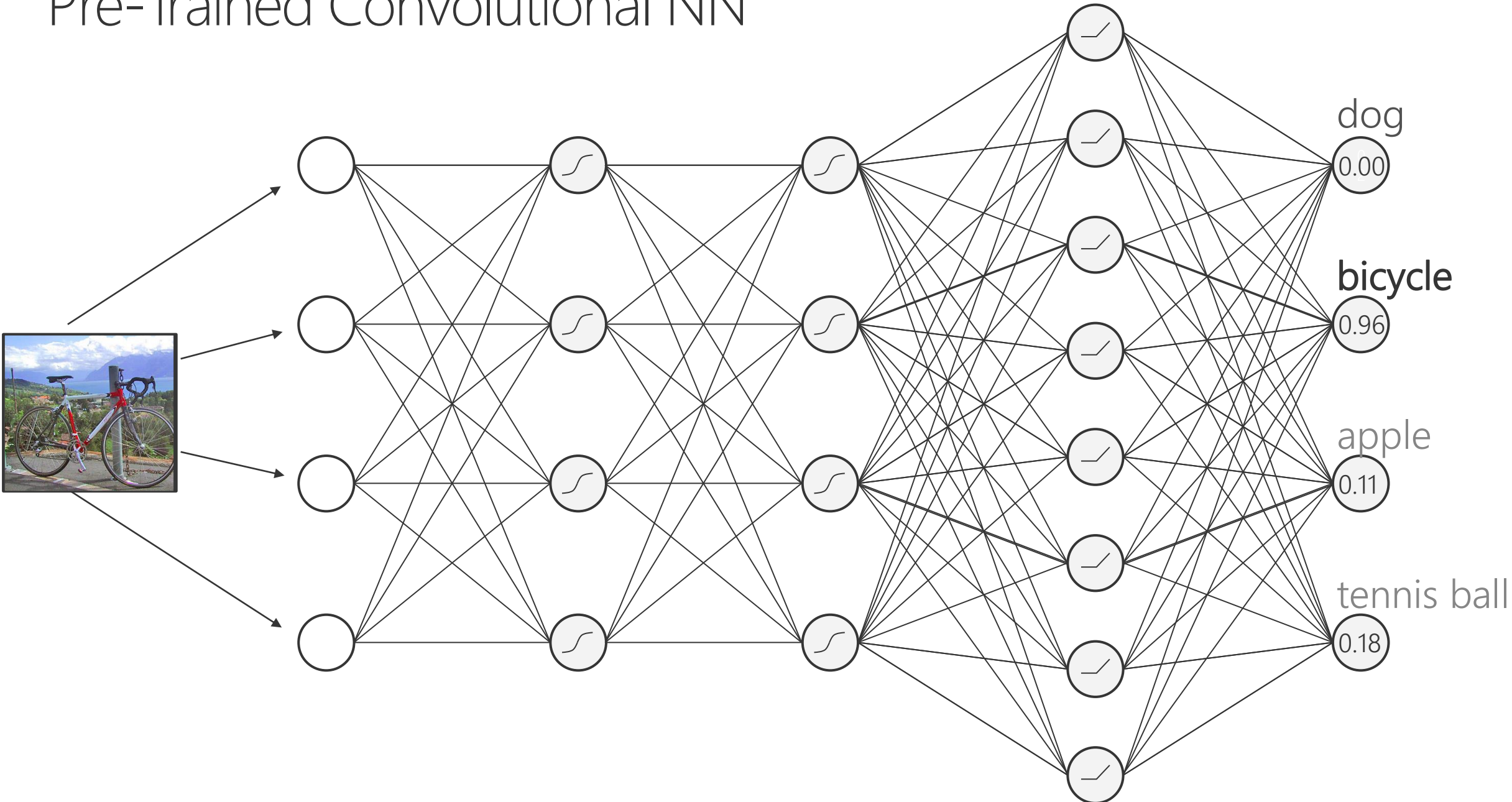
Convolutional Neural Network



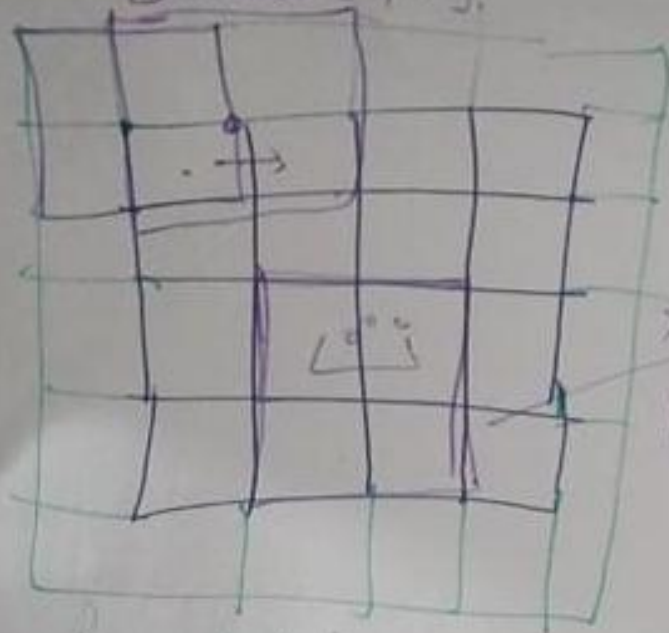
Trained Convolutional NN



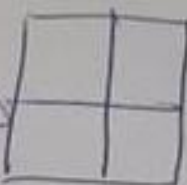
Pre-Trained Convolutional NN



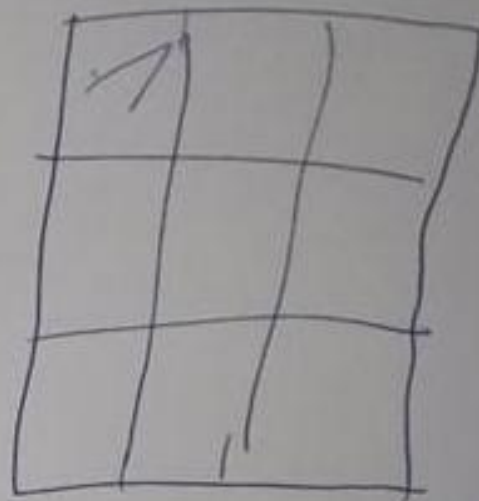
Stride (1,1)



filter



pooling



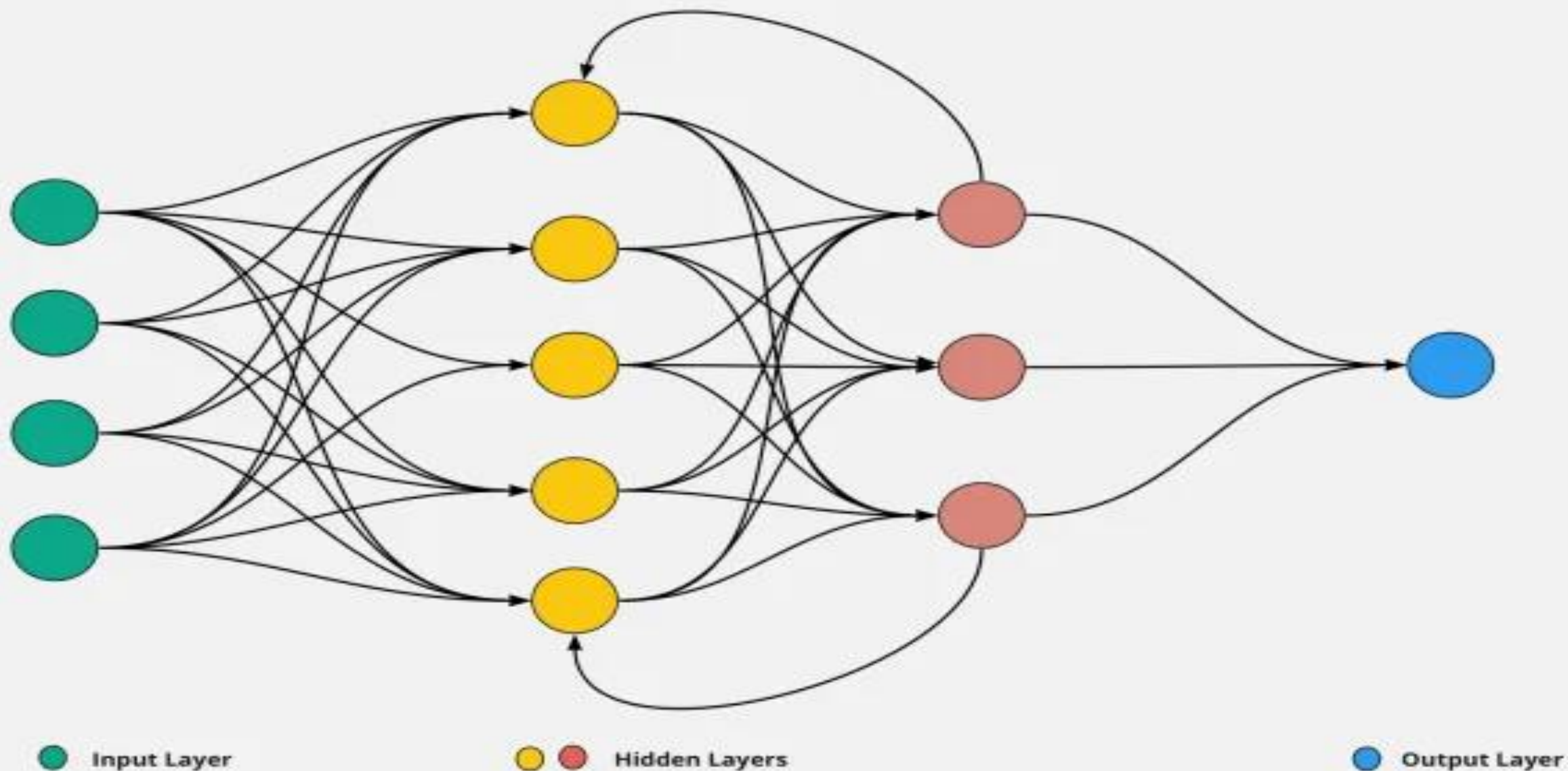
padding

elem must

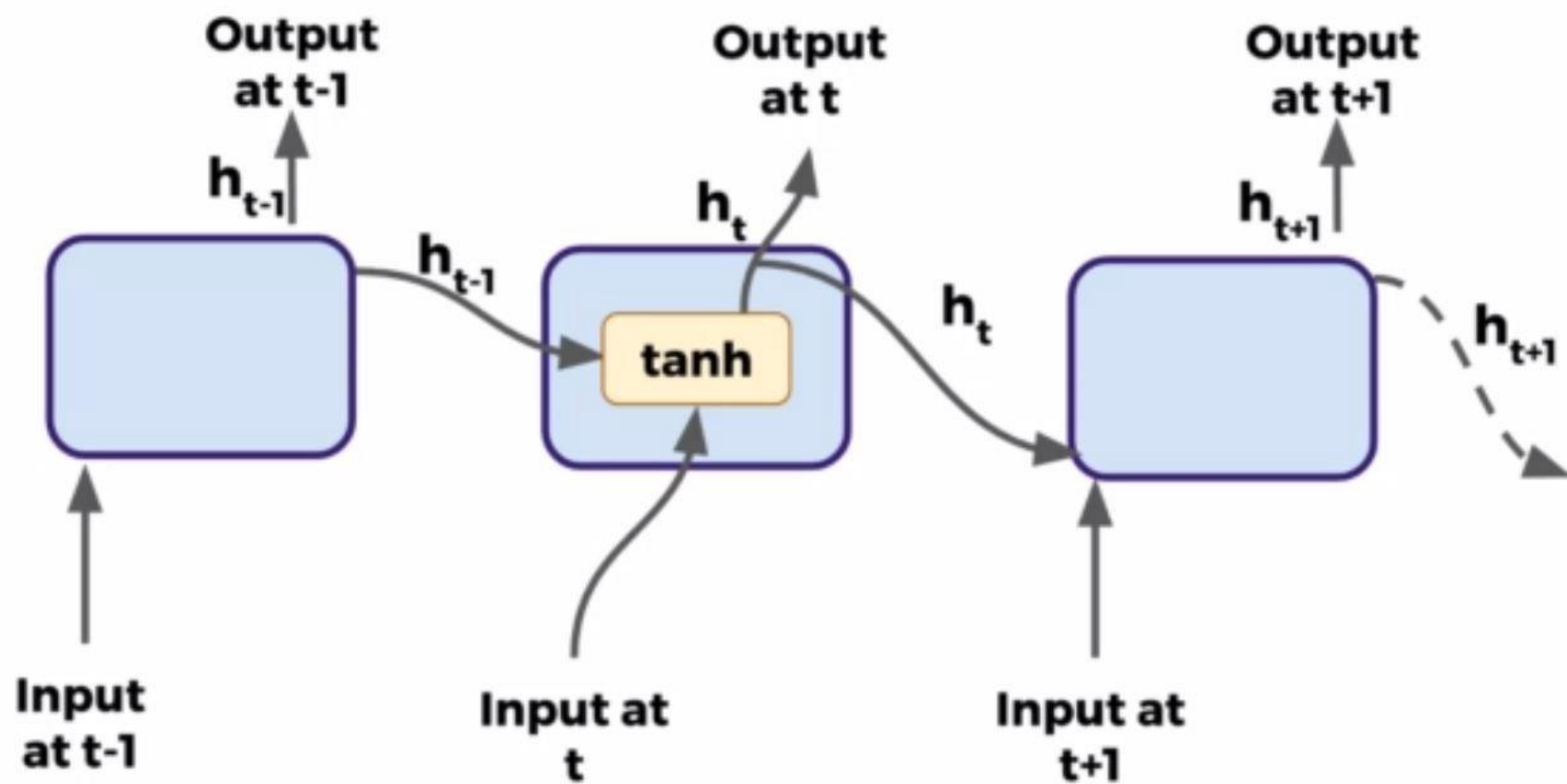
stride, pad, filter

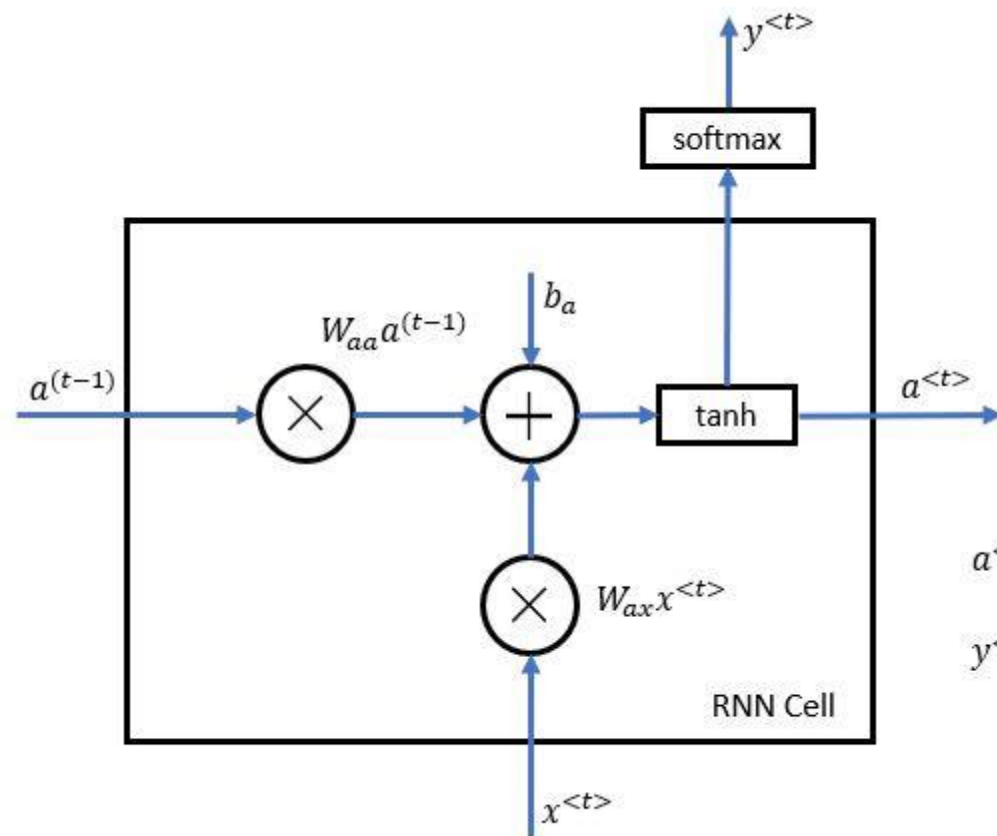
pooling

Recurrent Neural Network



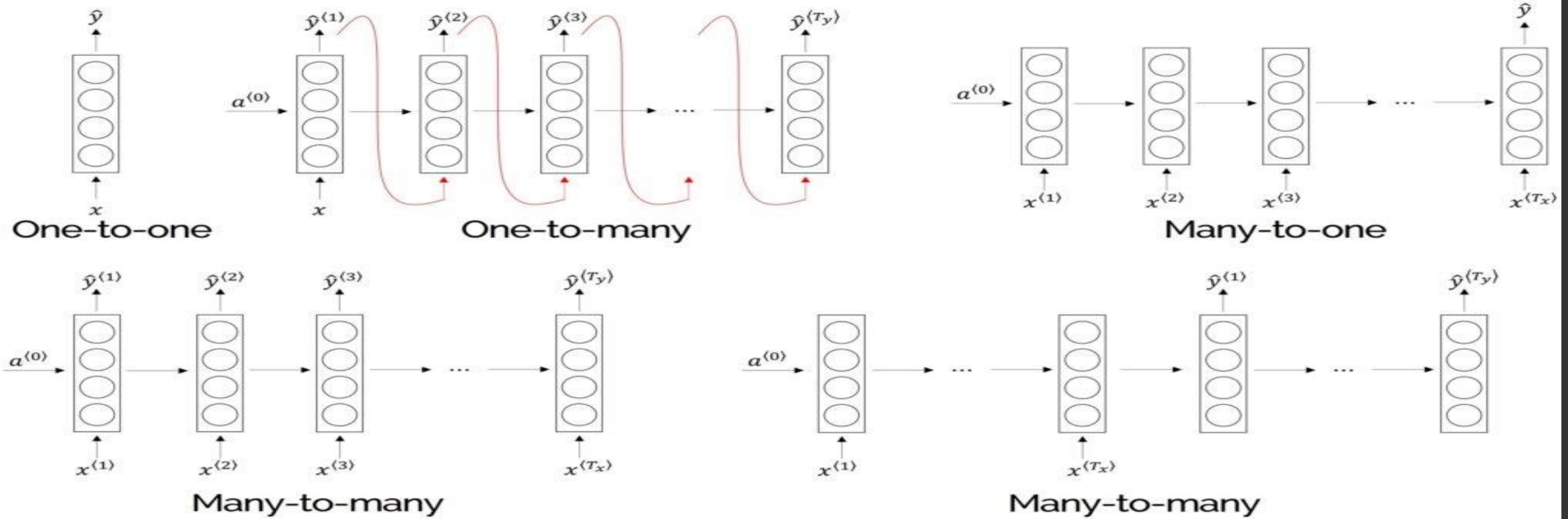
A typical RNN cell



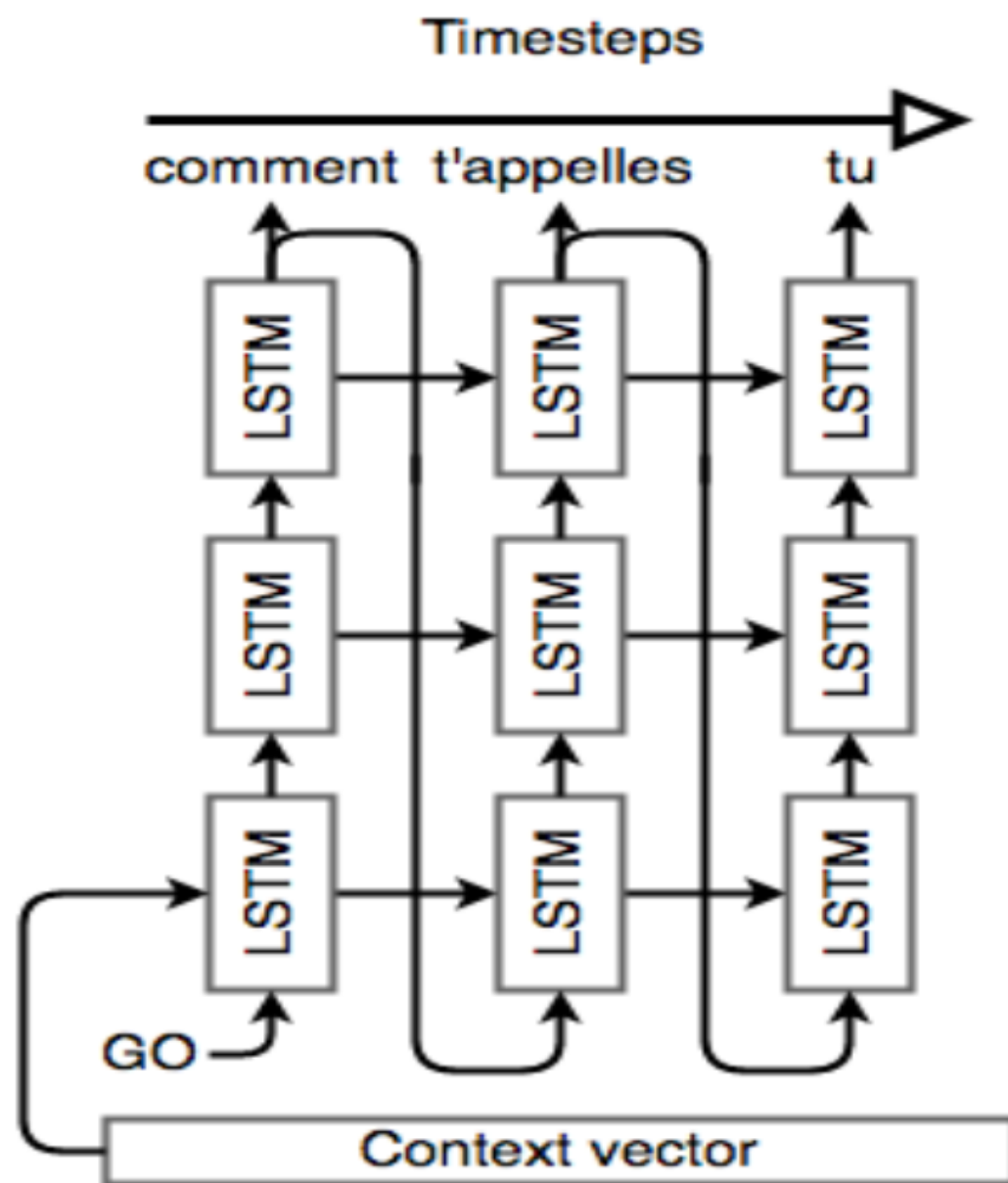


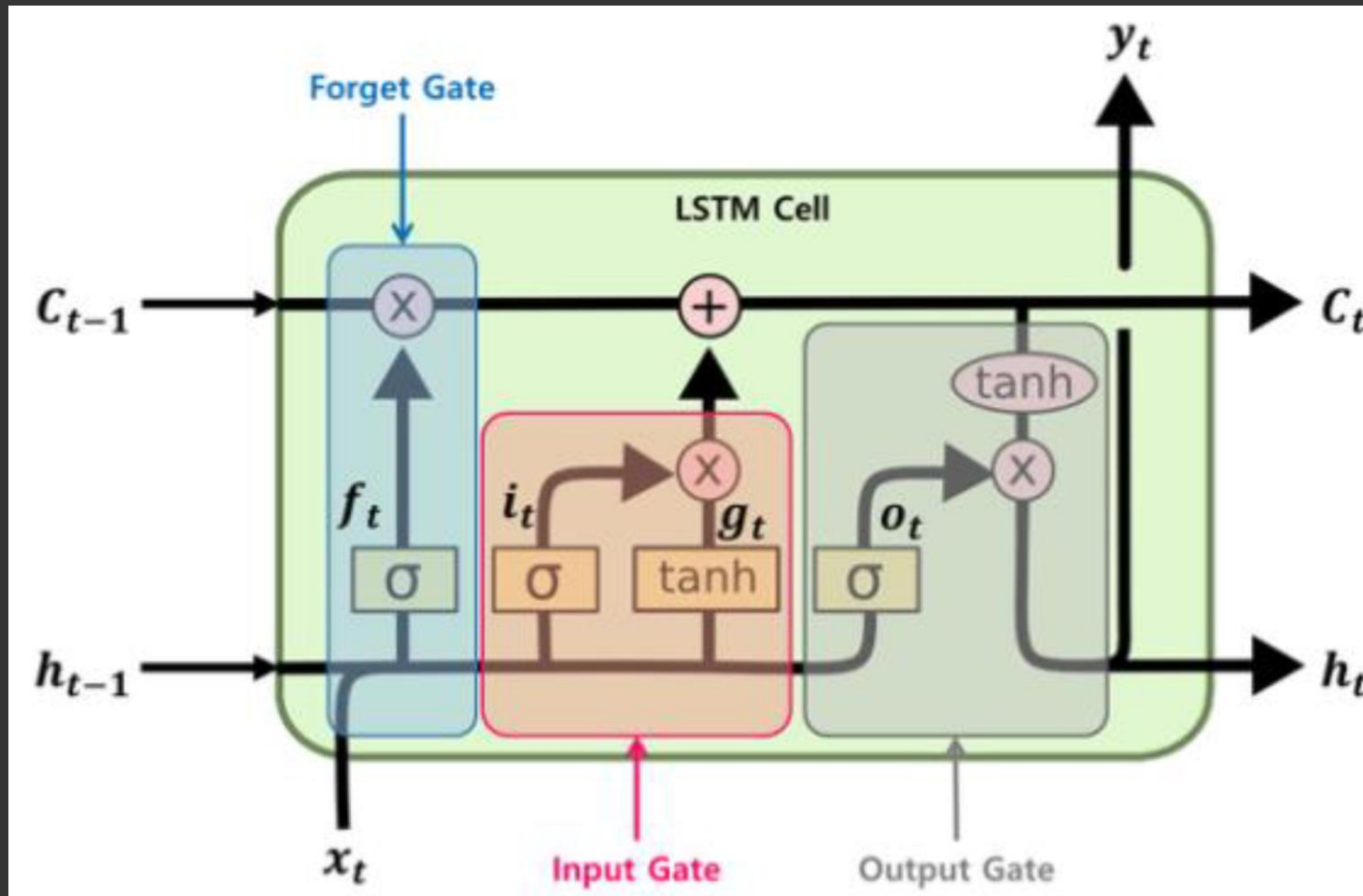
$$a^{<t>} = \tanh(W_{ax}x^{<t>} + W_{aa}a^{(t-1)} + b_a)$$

$$y^{<t>} = \text{softmax}(W_{ya}a^{<t>} + b_y)$$

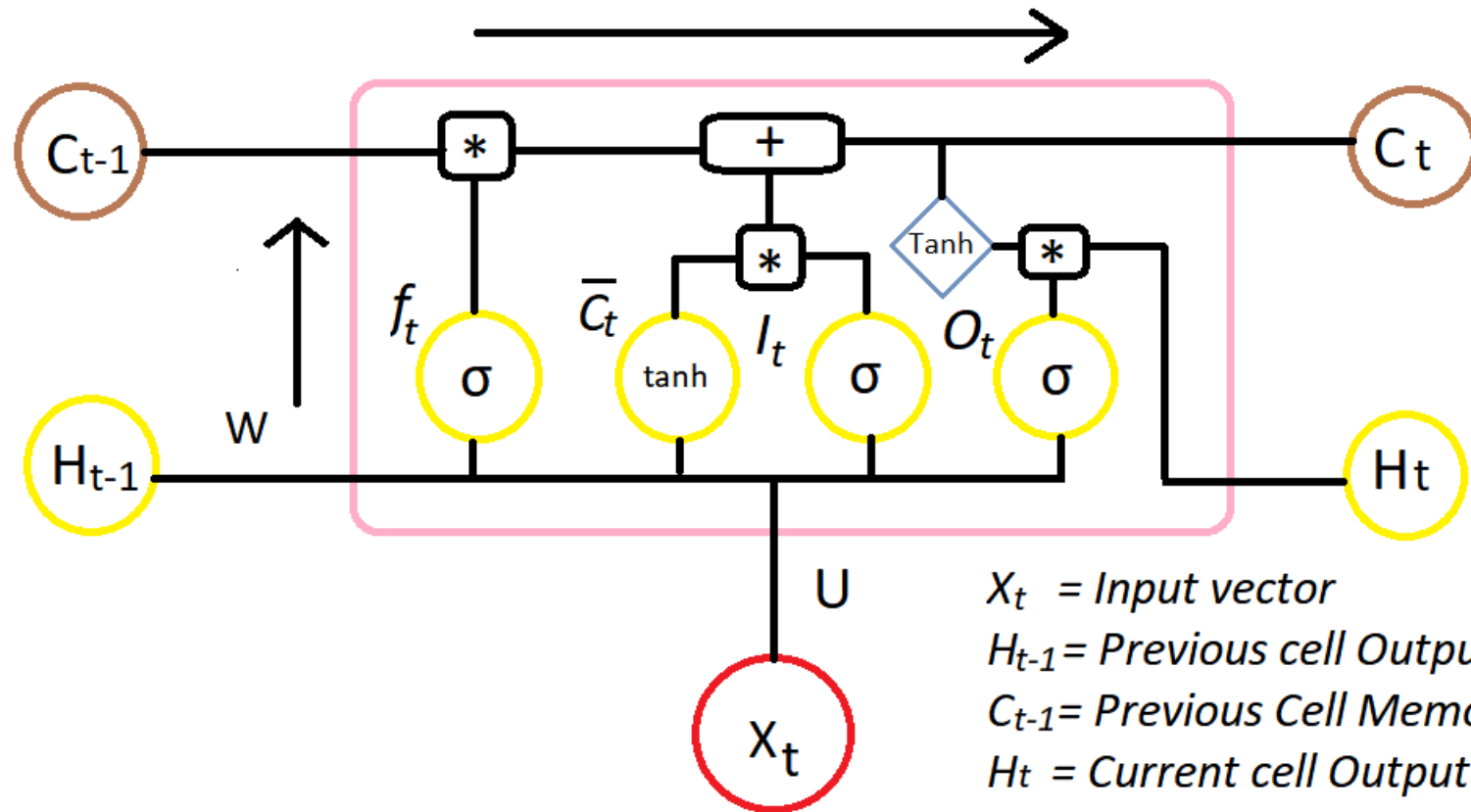


Architectural Types of Different Recurrent Neural Networks





LSTM Network



$*$ = Element-wise multiplication

$+$ = Element-wise addition

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

$$\bar{C}_t = \tanh (X_t * U_c + H_{t-1} * W_c)$$

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

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

$$C_t = f_t * C_{t-1} + I_t * \bar{C}_t$$

$$H_t = O_t * \tanh (C_t)$$

X_t = Input vector

H_{t-1} = Previous cell Output

C_{t-1} = Previous Cell Memory

H_t = Current cell Output

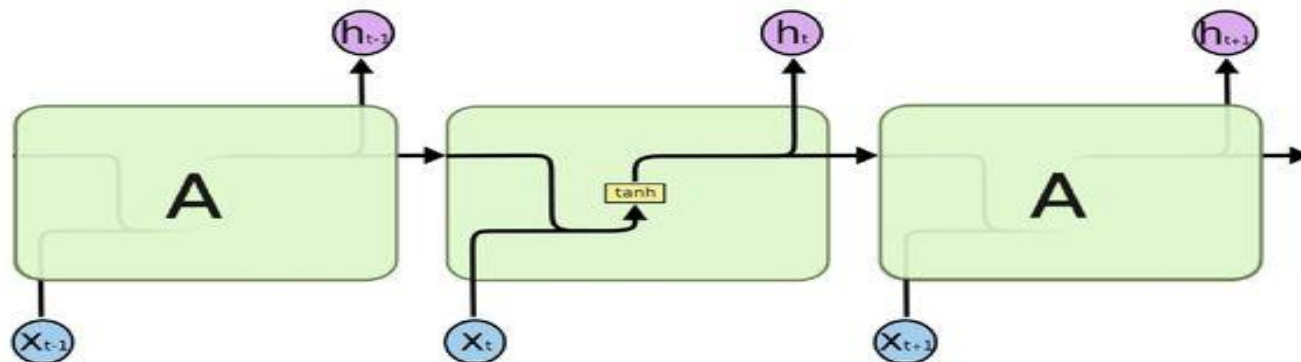
C_t = Current cell Memory

W, U = weight vectors for forget gate (f), candidate (c), i/p gate (i) and o/p gate (o)

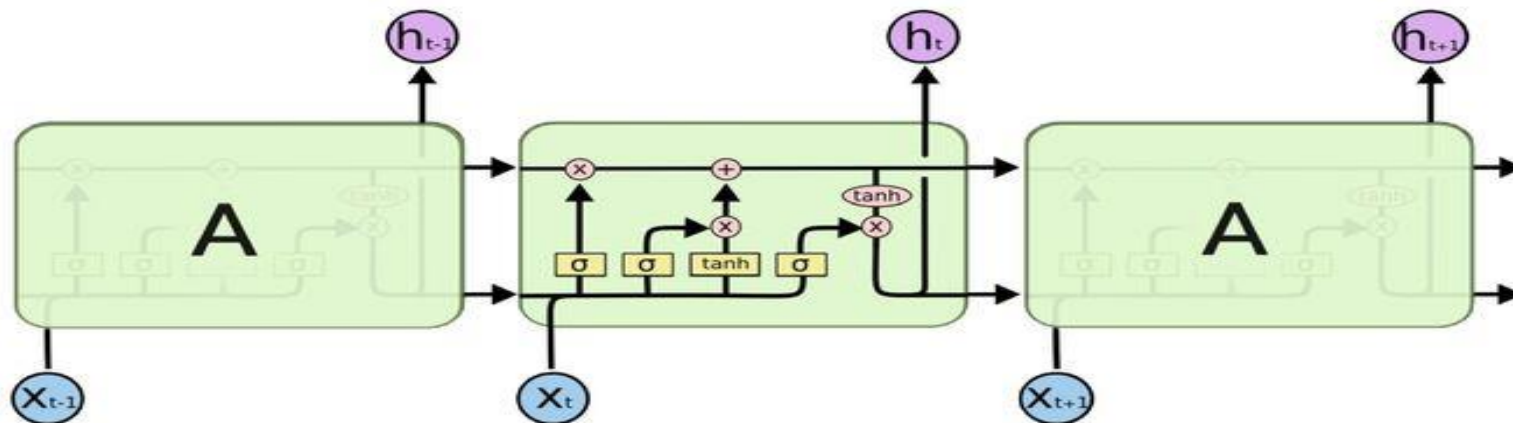
Note : These are different weights for different gates, for simplicity's sake, I mentioned W and U


LSTM (Long short-term memory)

- Standard RNN
- Input concatenate with output then feed to input again




- LSTM
- The repeating structure is more complicated





Neural Network
Layer


Pointwise
Operation


Vector
Transfer


Concatenate


Copy

Thank you!

Q & A