Deep Learning

By MXK

Summary

1- Machine Learning Roadmap

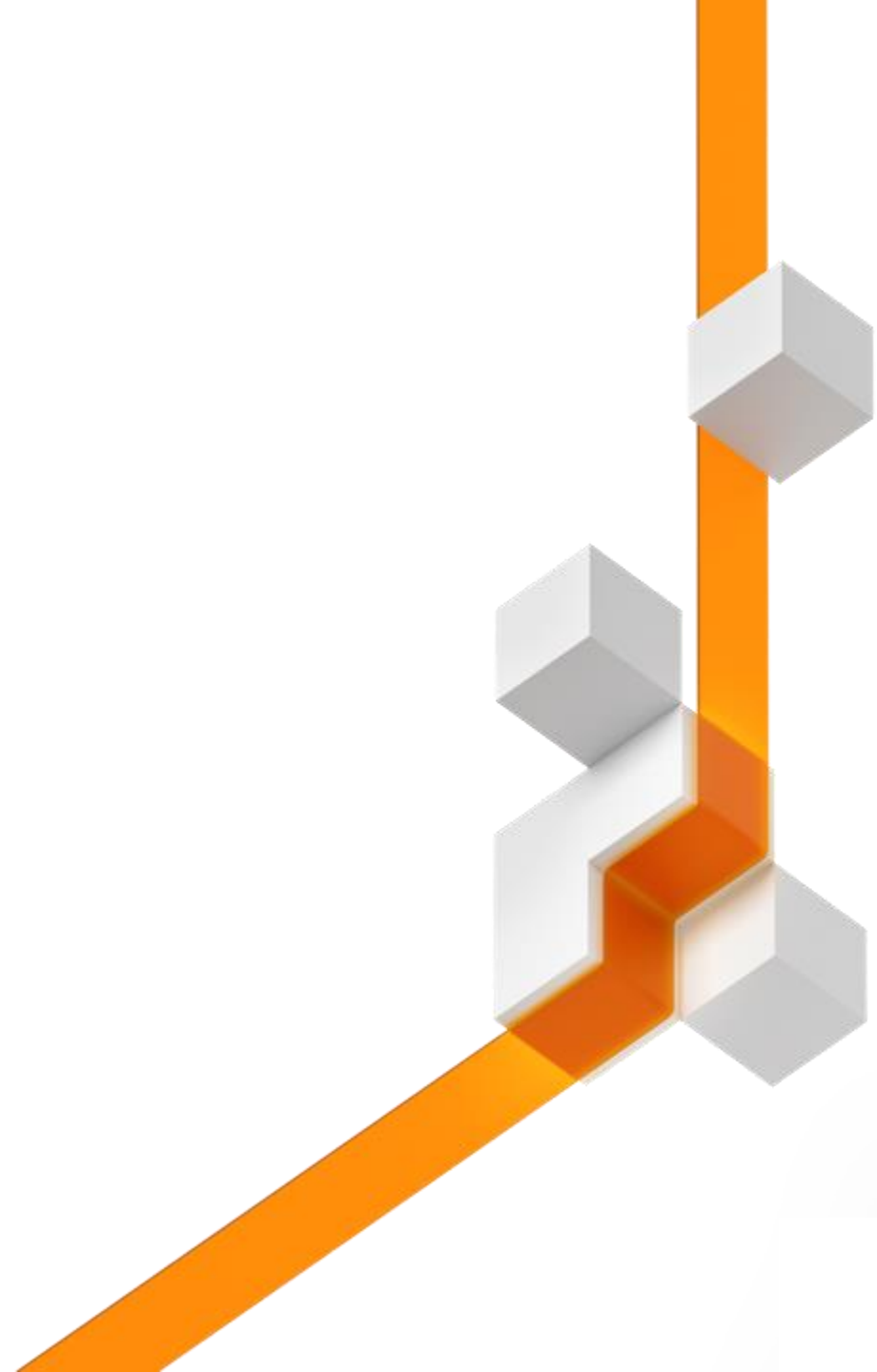
2- Deep Learning Roadmap

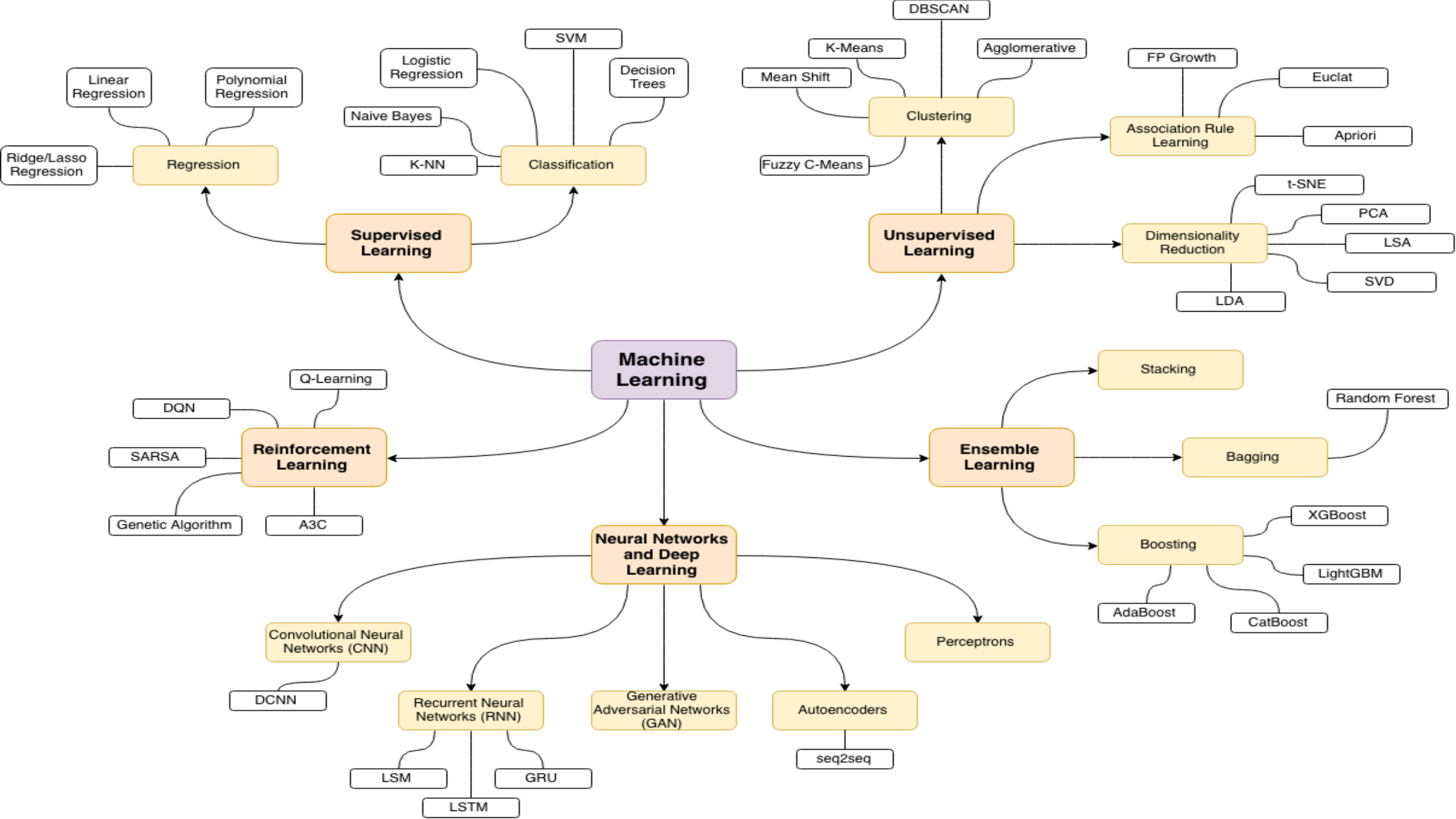
3- Convolutional Neural Network

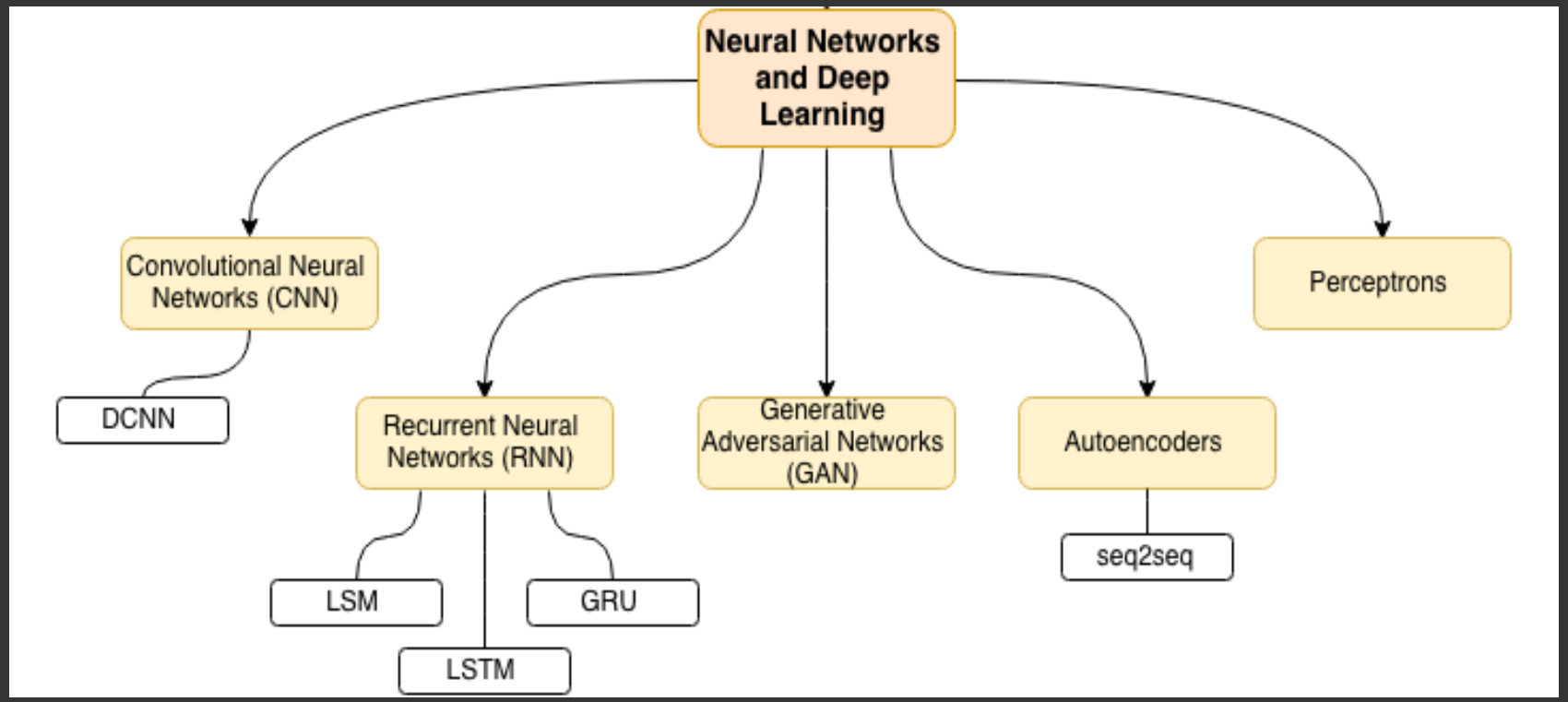
4- Recurrent Neural Network

5- Long Short Term Memory

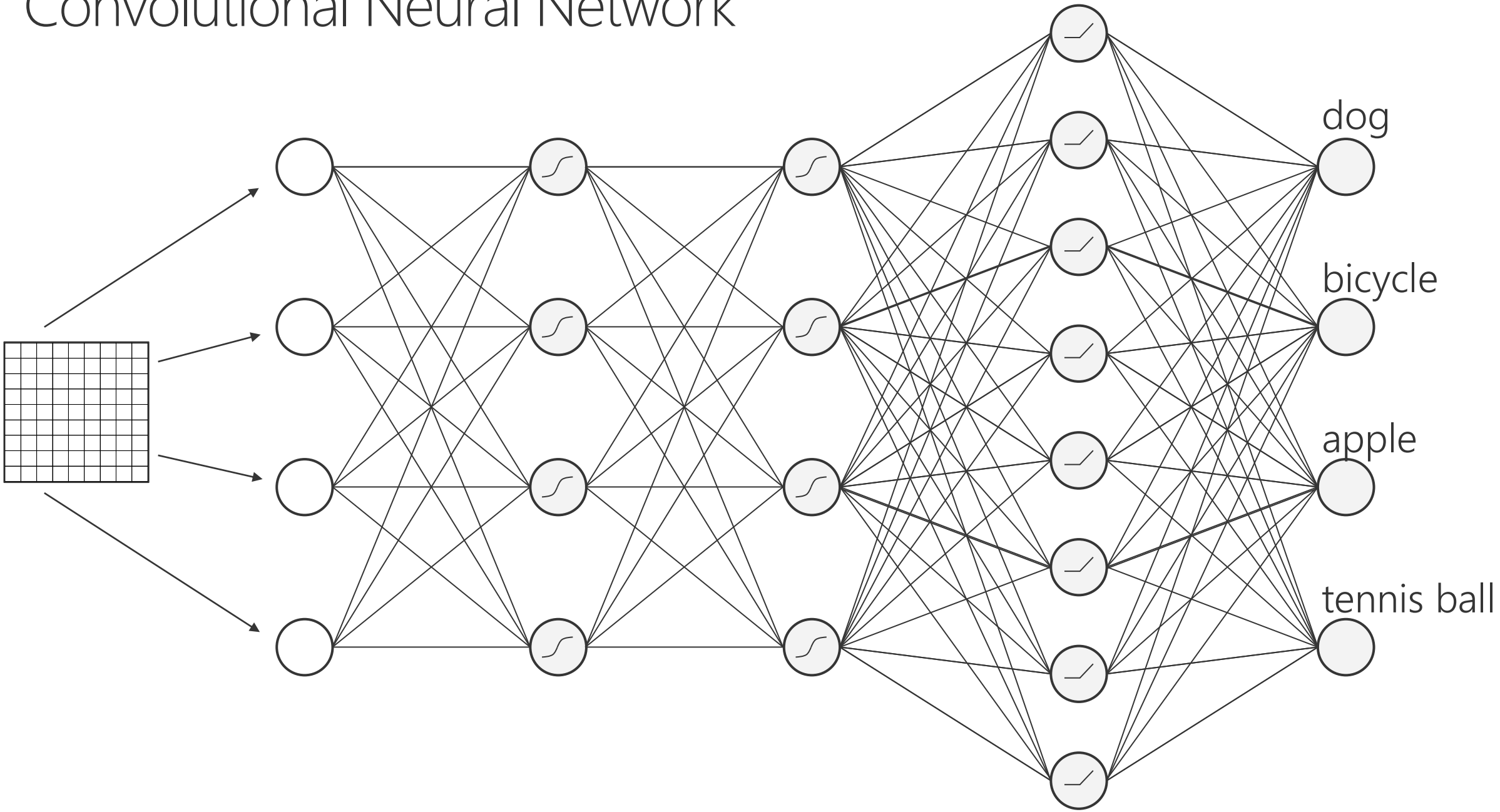
6- Closing Thoughts



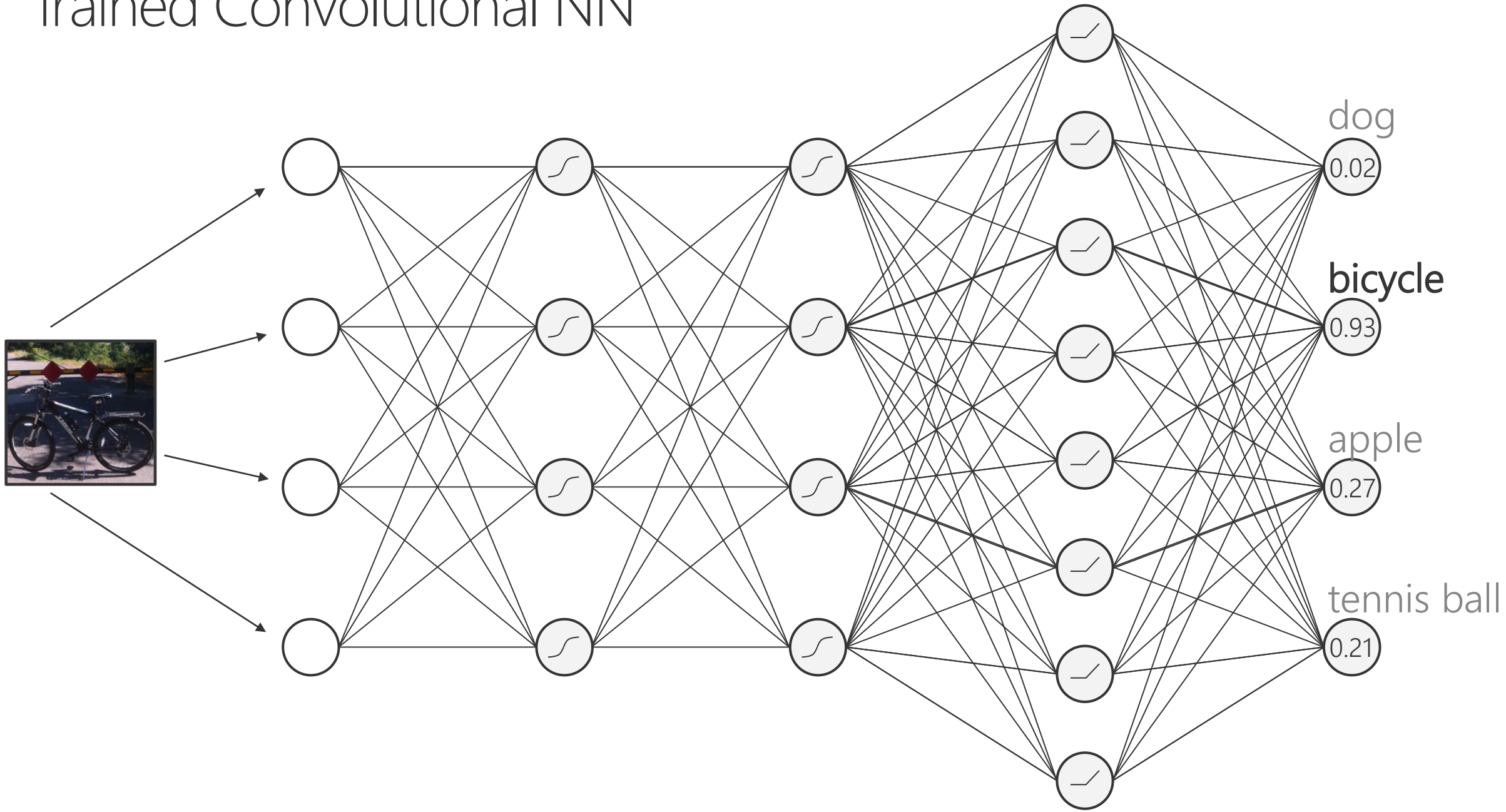

Machine Learning
Supervised Learning
Regression
Linear Regression
Polynomial Regression
Ridge/Lasso Regression
Classification
SVM
Logistic Regression
Decision Trees
Naive Bayes
K-NN
Unsupervised Learning
Clustering
DBSCAN
K-Means
Agglomerative
Mean Shift
Fuzzy C-Means
Association Rule Learning
FP Growth
Euclat
Apriori
Dimensionality Reduction
t-SNE
PCA
LSA
SVD
LDA
Reinforcement Learning
Q-Learning
DQN
SARSA
Genetic Algorithm
A3C
Ensemble Learning
Stacking
Bagging
Random Forest
Boosting
XGBoost
LightGBM
AdaBoost
CatBoost
Neural Networks and Deep Learning
Convolutional Neural Networks (CNN)
DCNN
Recurrent Neural Networks (RNN)
LSM
GRU
LSTM
Generative Adversarial Networks (GAN)
Autoencoders
seq2seq
Perceptrons


Neural Networks and Deep Learning
Convolutional Neural Networks (CNN)
DCNN
Recurrent Neural Networks (RNN)
LSM
GRU
LSTM
Generative Adversarial Networks (GAN)
Autoencoders
seq2seq
Perceptrons

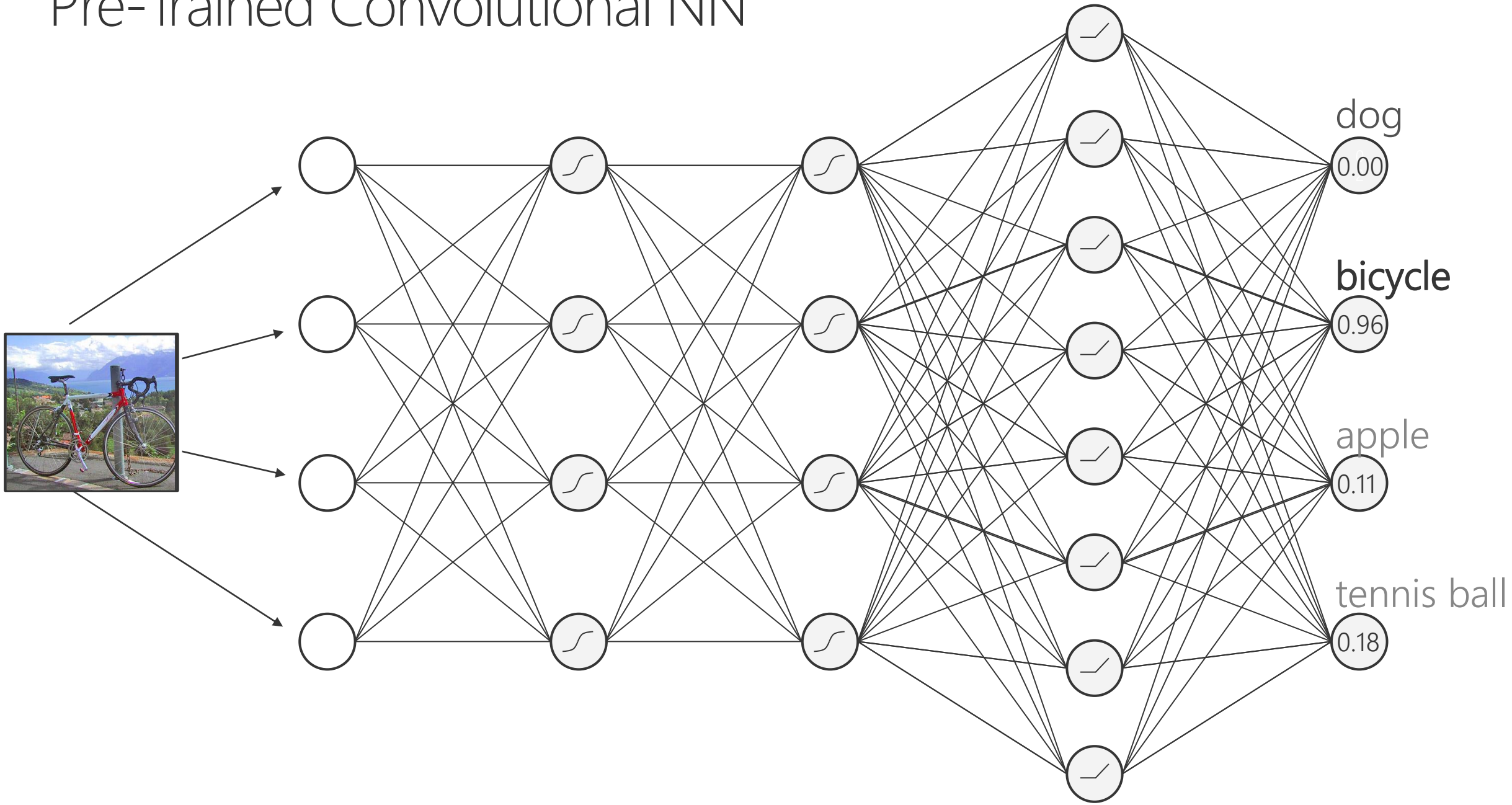
Convolutional Neural Network



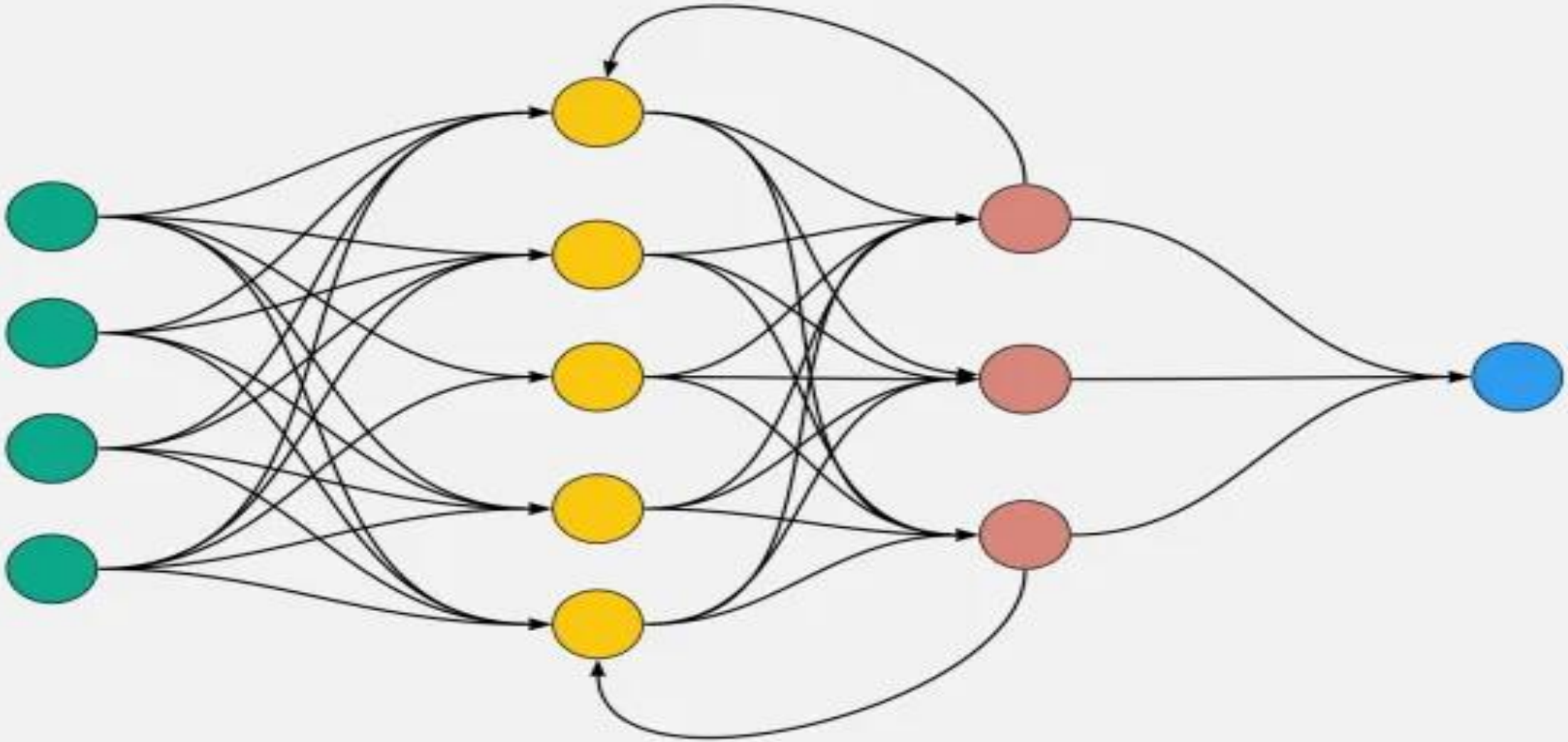
Trained Convolutional NN



Pre-Trained Convolutional NN



Recurrent Neural Network



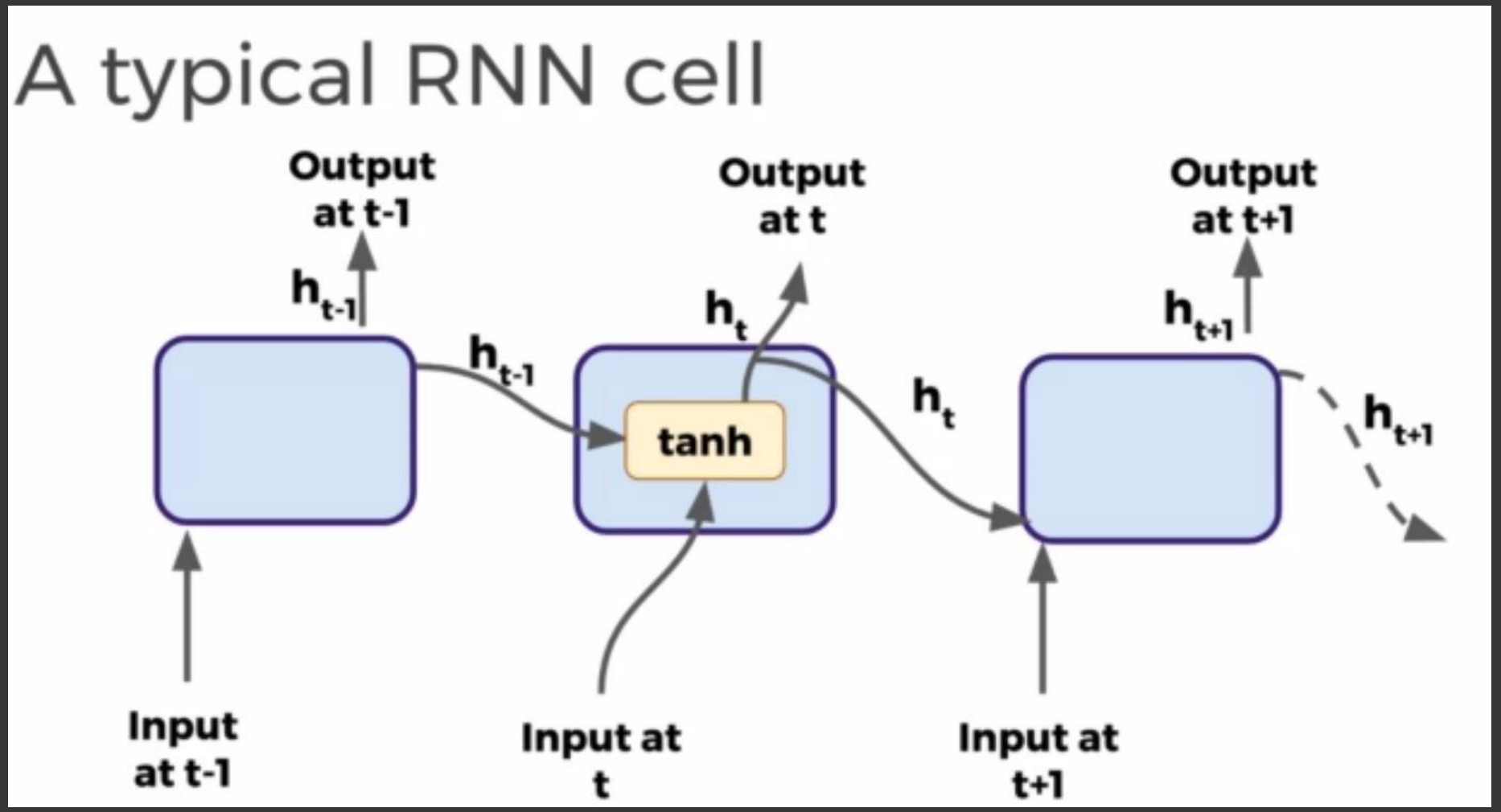
Input Layer

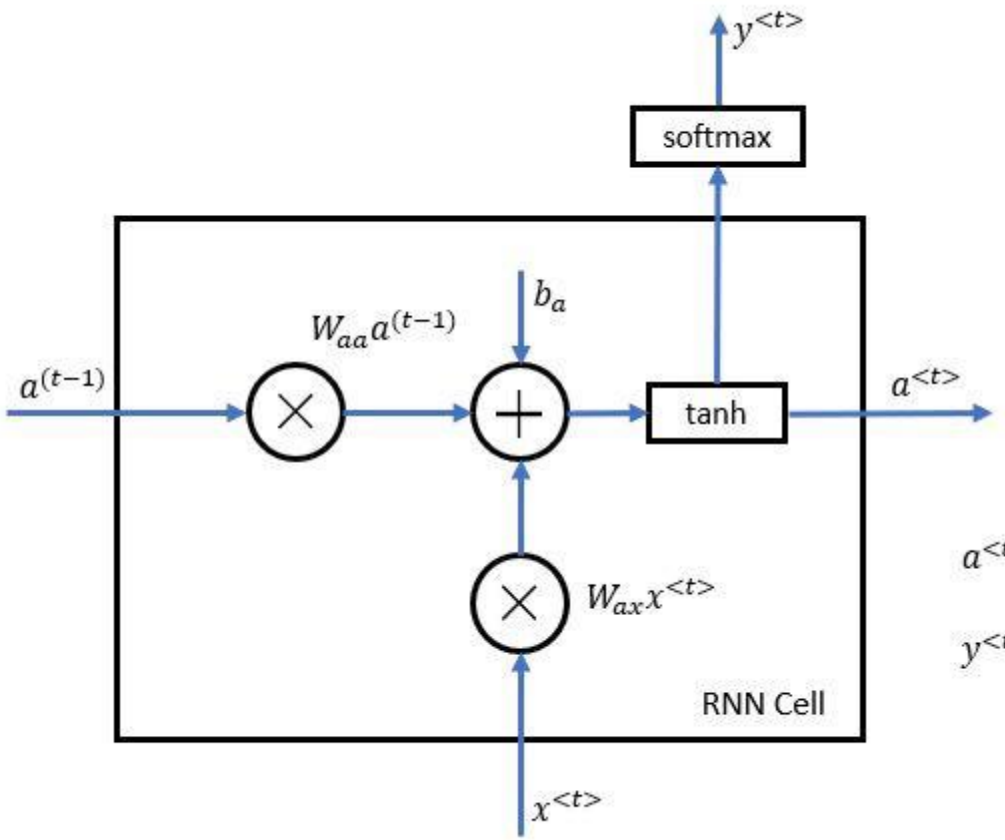
Hidden Layers

Output Layer

dataaspirant.com

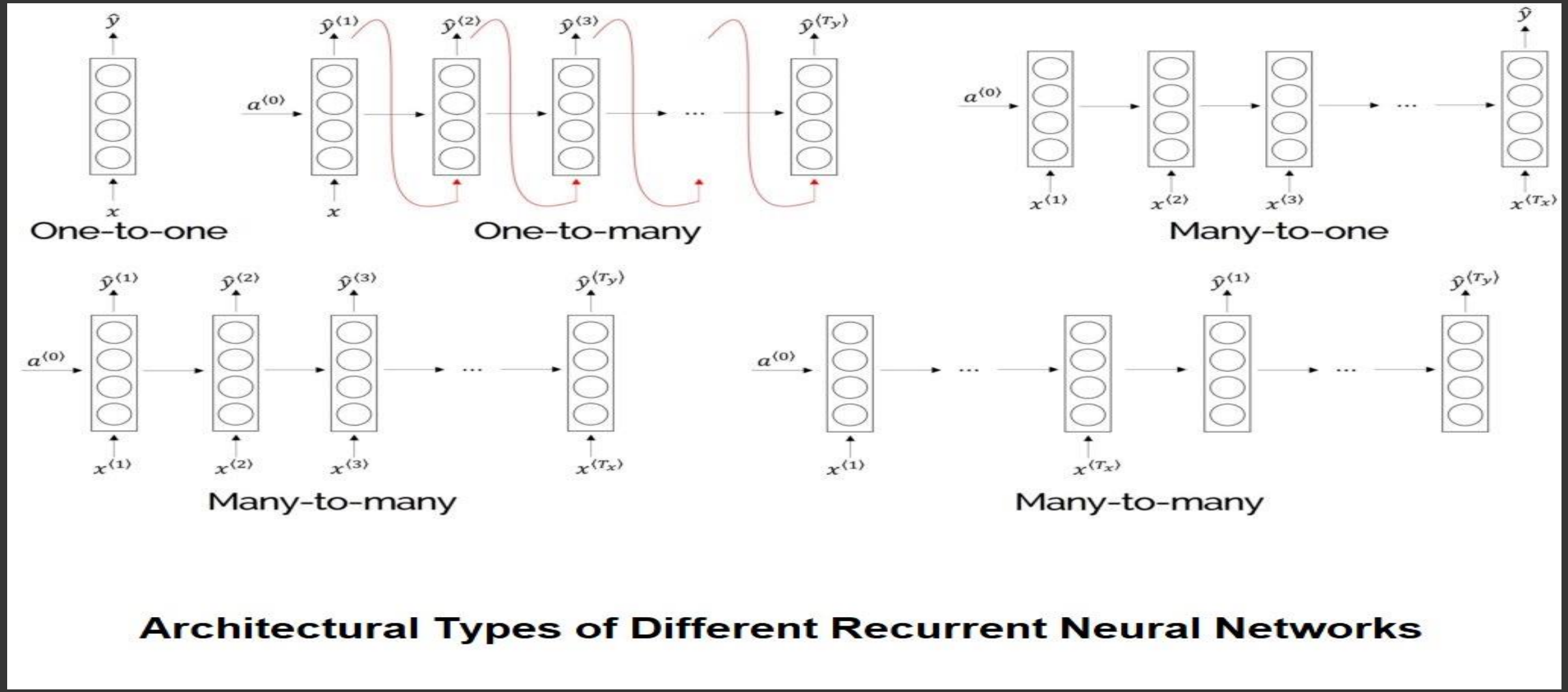
A typical RNN cell



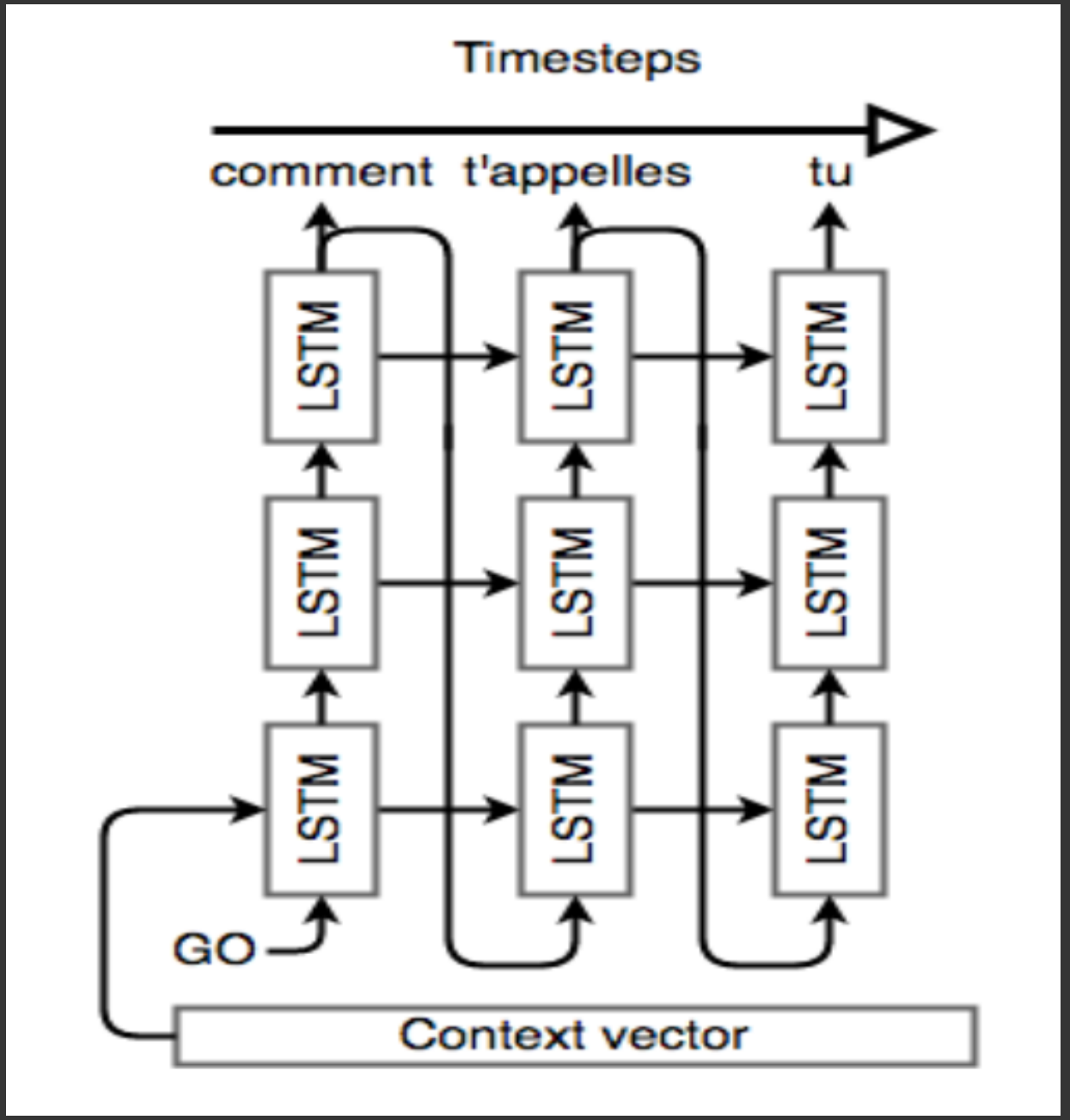


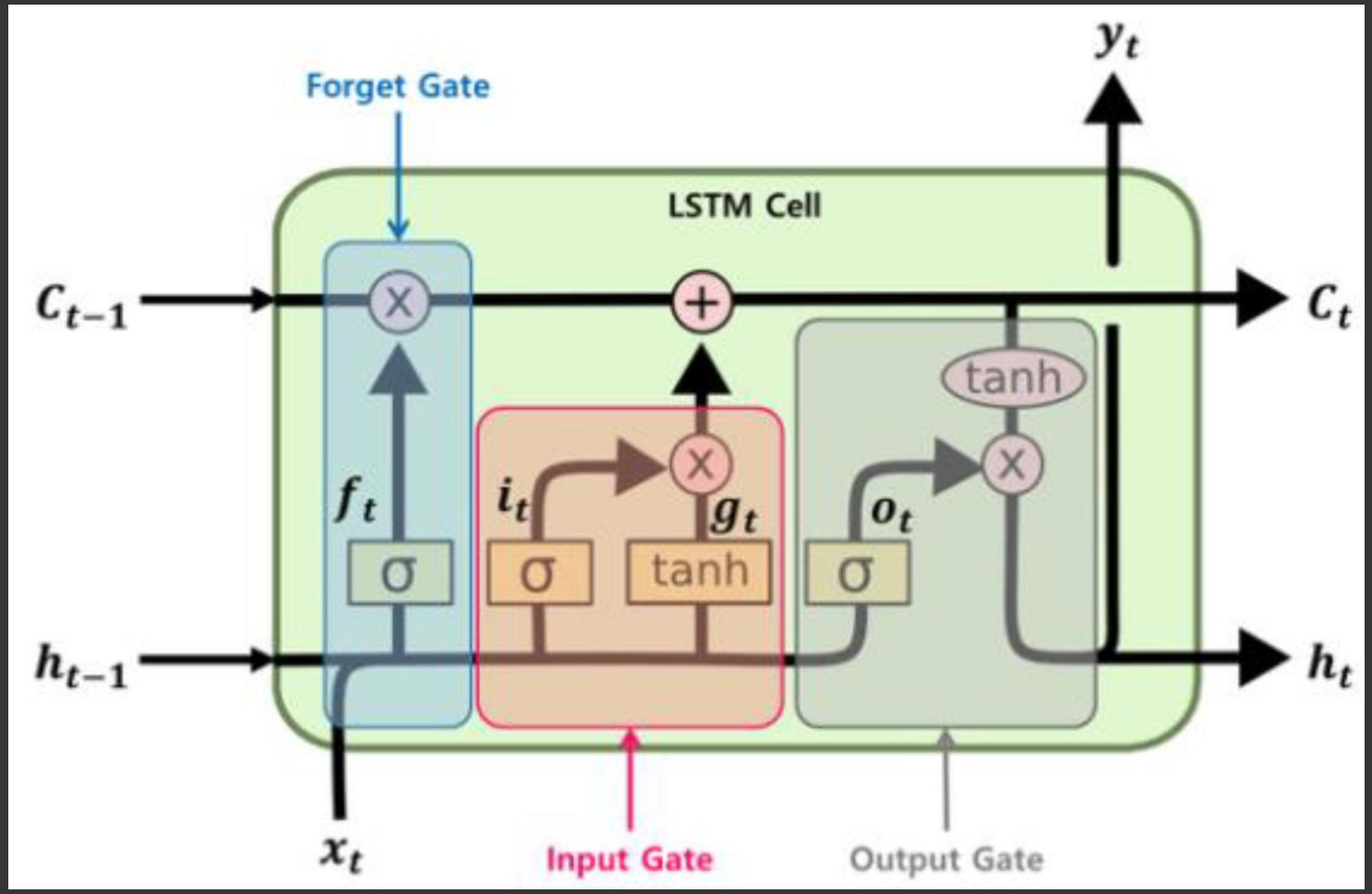
$$a^{<t>} = \tanh(W_{ax}x^{<t>} + W_{aa}a^{(t-1)} + b_a)$$

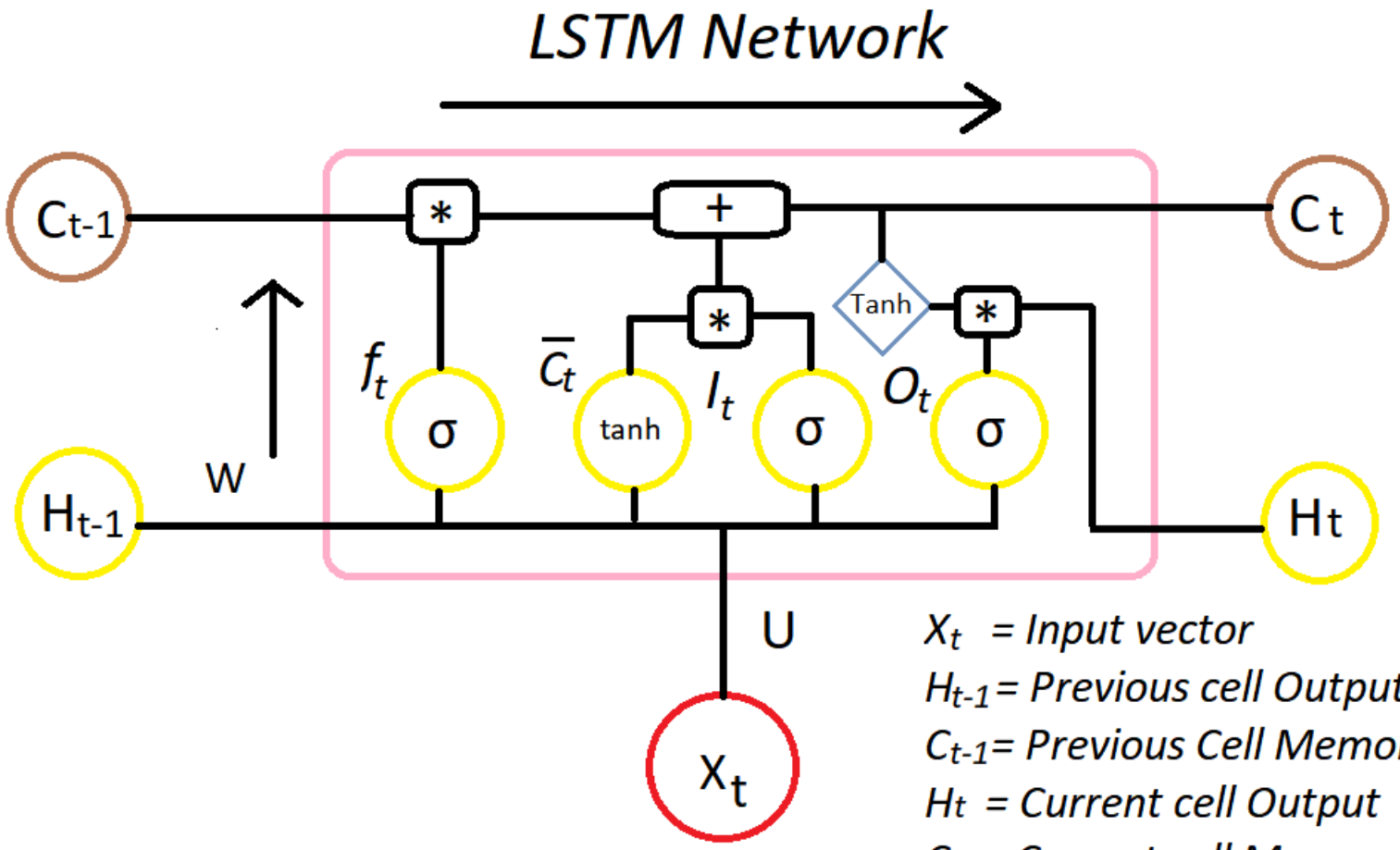
$$y^{<t>} = softmax(W_{ya}a^{<t>} + b_y)$$



Architectural Types of Different Recurrent Neural Networks


Timesteps
comment
t'appelles
tu
LSTM
LSTM
LSTM
LSTM
LSTM
LSTM
LSTM
LSTM
LSTM
GO
Context vector


y_t
Forget Gate
LSTM Cell
C_{t-1}
C_t
tanh
f_t
i_t
g_t
o_t
σ
σ
tanh
σ
h_{t-1}
h_t
x_t
Input Gate
Output Gate



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

(*) = Element-wise multiplication

(+) = Element-wise addition

$$f_t = \sigma (X_t * U_f + H_{t-1} * W_f)$$
$$\bar{C}_t = \tanh (X_t * U_c + H_{t-1} * W_c)$$
$$I_t = \sigma (X_t * U_i + H_{t-1} * W_i)$$
$$O_t = \sigma (X_t * U_o + H_{t-1} * W_o)$$

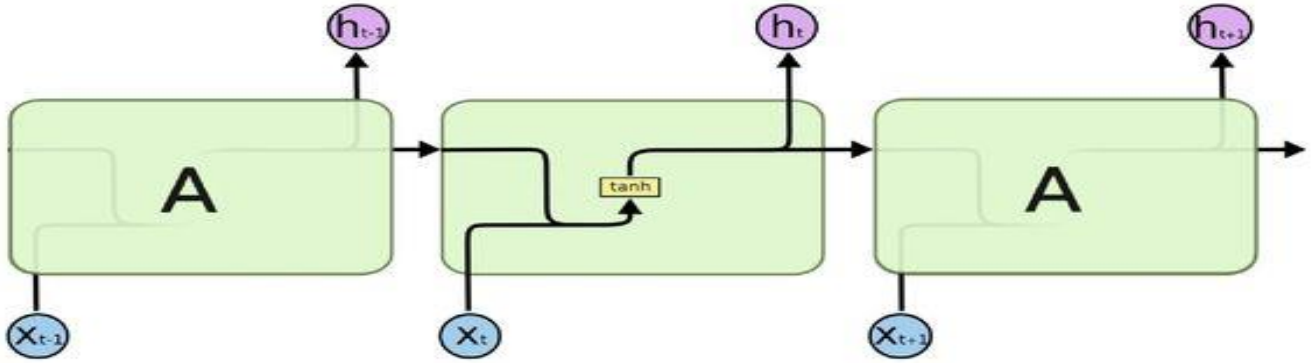
$$C_t = f_t * C_{t-1} + I_t * \bar{C}_t$$
$$H_t = O_t * \tanh (C_t)$$

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

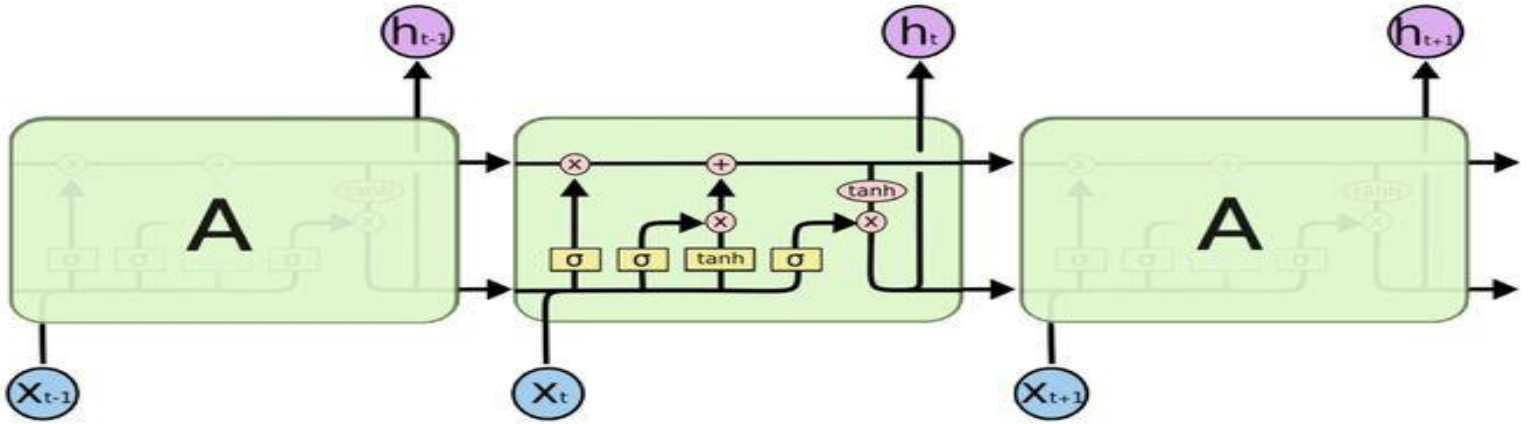
Note : These are different weights for different gates, for simpicity's sake, I mentioned W and U

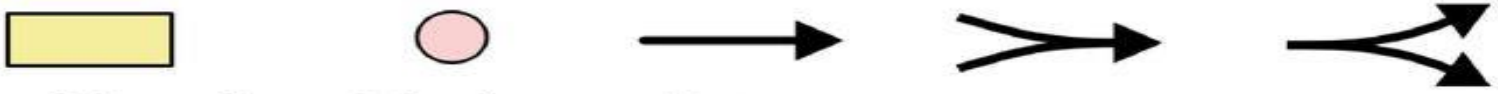
LSTM (Long short-term memory)

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



- LSTM
- The repeating structure is more complicated





Thank you!

Q & A