Recurrent Neural Network

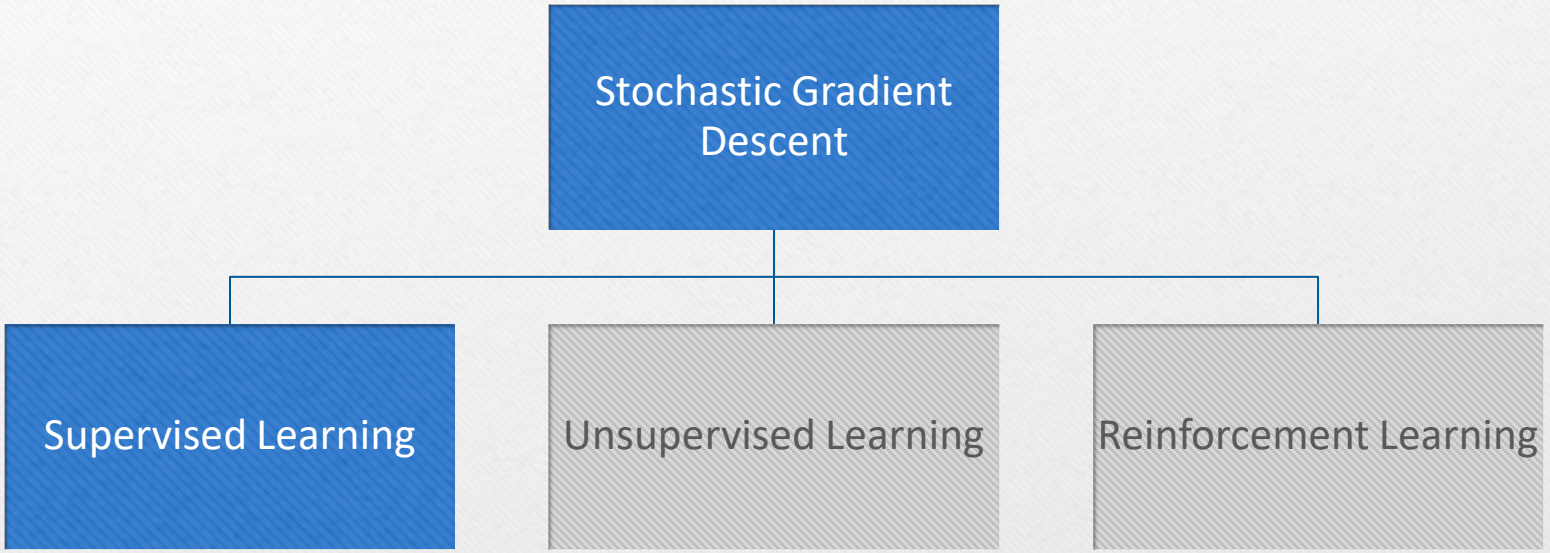
ECON 4810

Important ANN Architectures

Plus Important Specific Applications

- Fully-Connected Neural Network (**MLP**)
  - Recurrent Neural Network (**RNN**)
    - Long Short-Term Memory (**LSTM**)
      - Bidirectional RNN
        - ELMo
  - Convolutional Neural Network (**CNN**)
    - AlexNet
    - Residual Learning (**ResNet**)
      - AlphaGo Zero
      - WaveNet
  - Transformer
    - BERT
    - GPT
    - AlphaStar

How ANN Learns



Recurrent Neural Network

RNN has autoregressive layers.

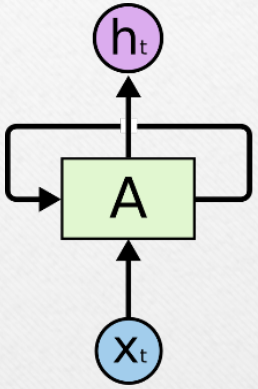


Image Source: colah's blog

Recurrent Neural Network

RNN has autoregressive layers.

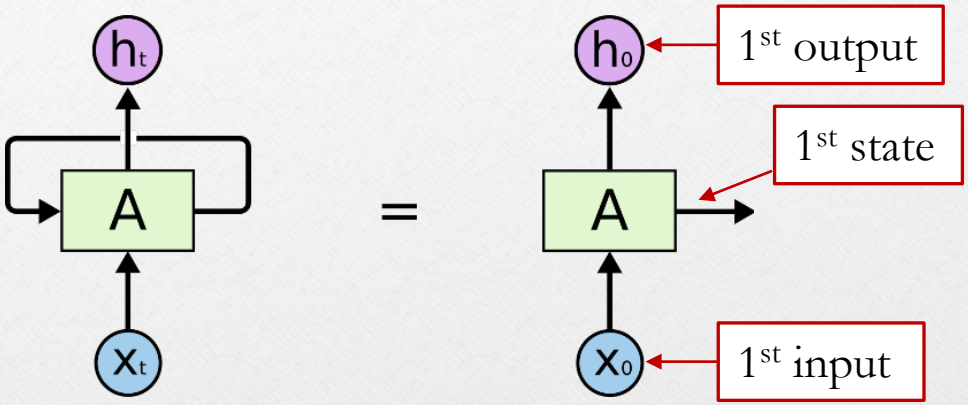


Image Source: colah's blog

Recurrent Neural Network

RNN has autoregressive layers.

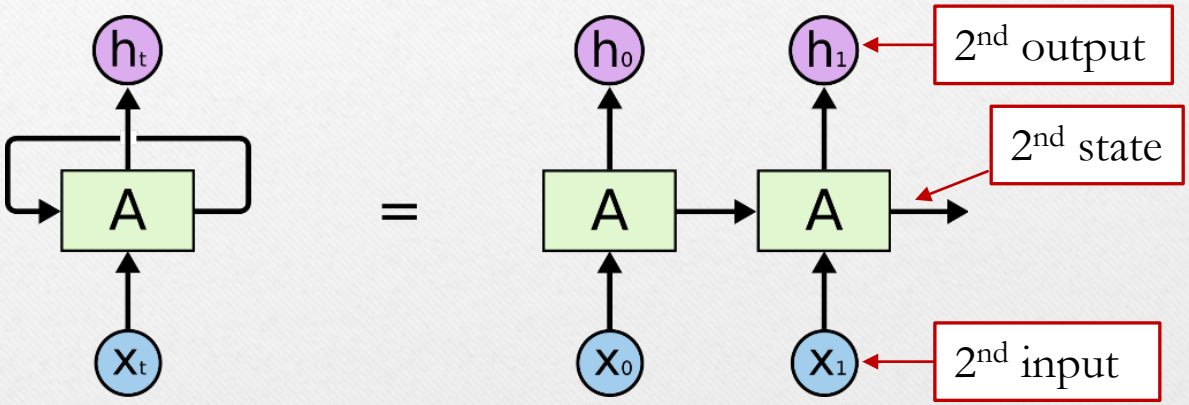


Image Source: colah's blog

Recurrent Neural Network

RNN has autoregressive layers.

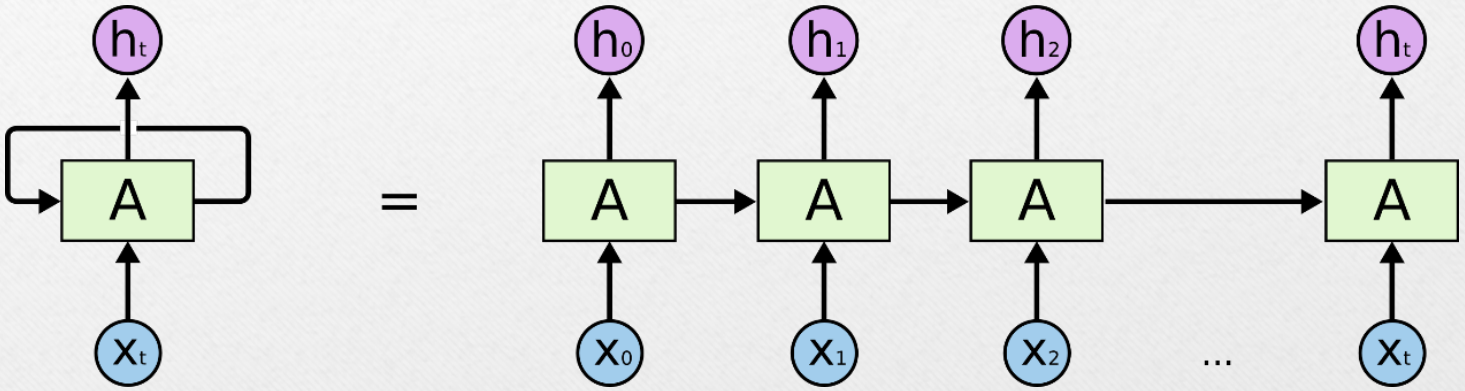
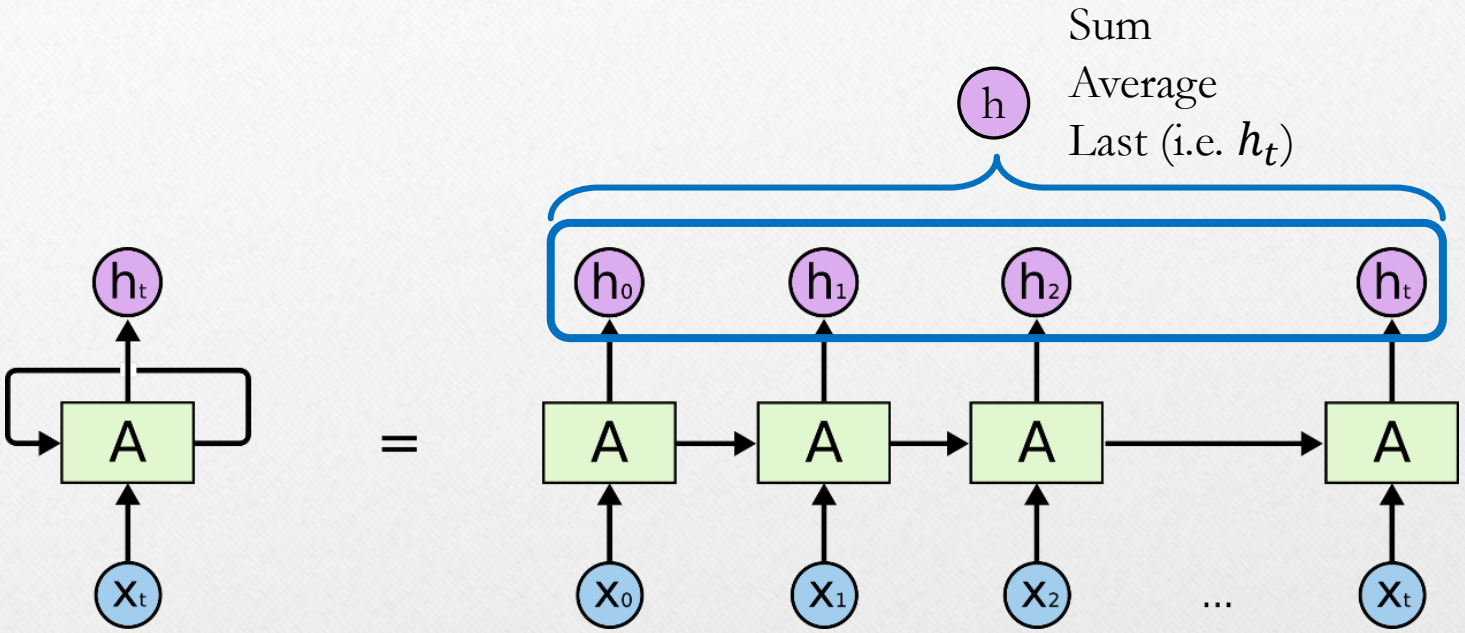


Image Source: colah's blog

Recurrent Neural Network

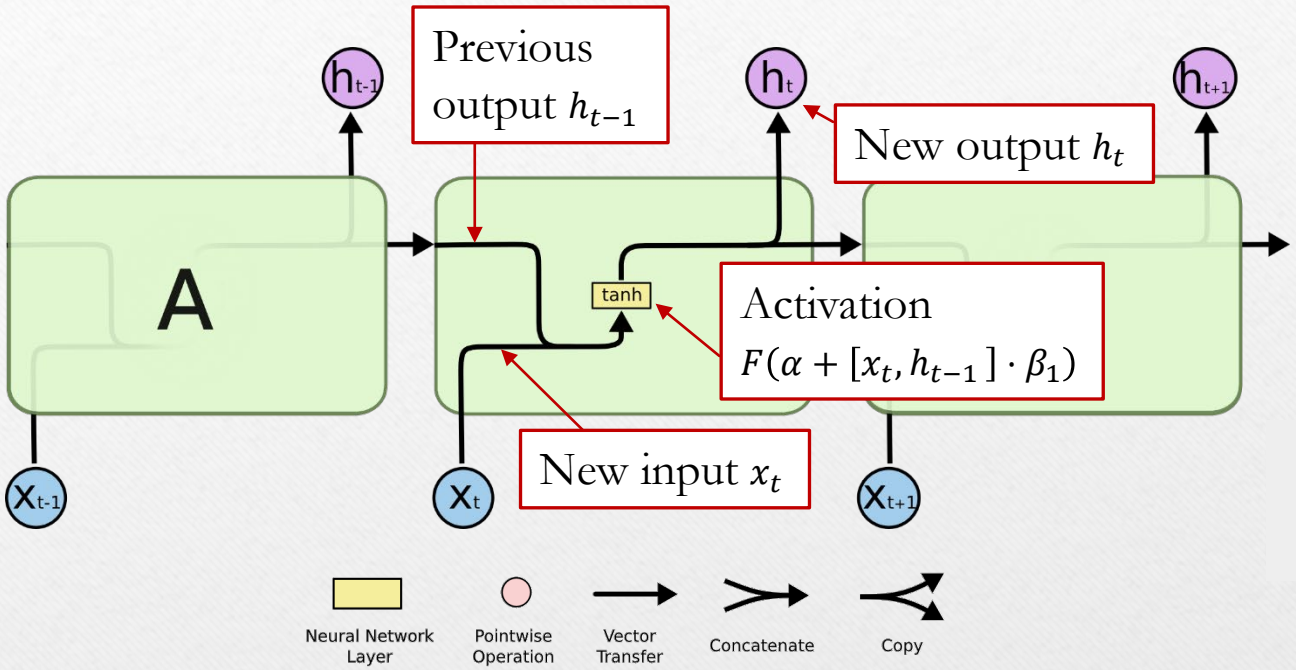


Word Level	How	Are	You	...	(Padding)
Character Level	H	O	W	...	(Padding)
Time Series	42.5	38	39.3	...	37

Image Source: colah's blog

Common Types of RNN

Standard RNN

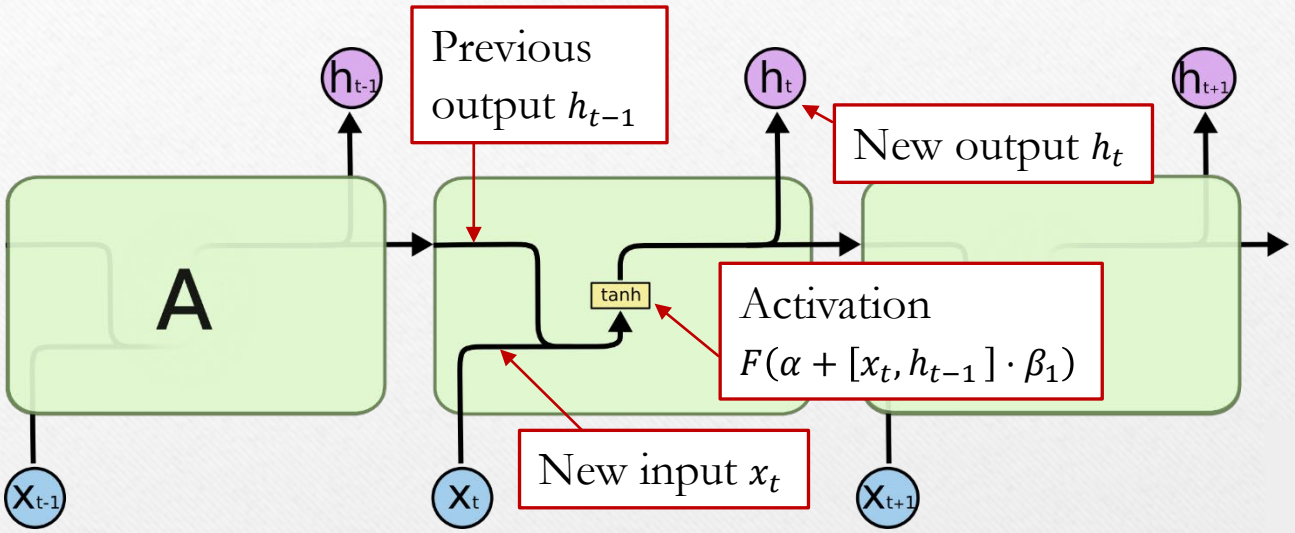


- Standard RNN applies the same weights to previous output in every step. It is incapable of assigning a higher weight to a particular time step.
- It is therefore generally incapable of "remembering" information from many time steps ago.
- When humans read, we do assign different weights to different information. We understand context and selectively remember things.

Image Source: colah's blog

Common Types of RNN

Standard RNN



Long Short Term Memory (LSTM)

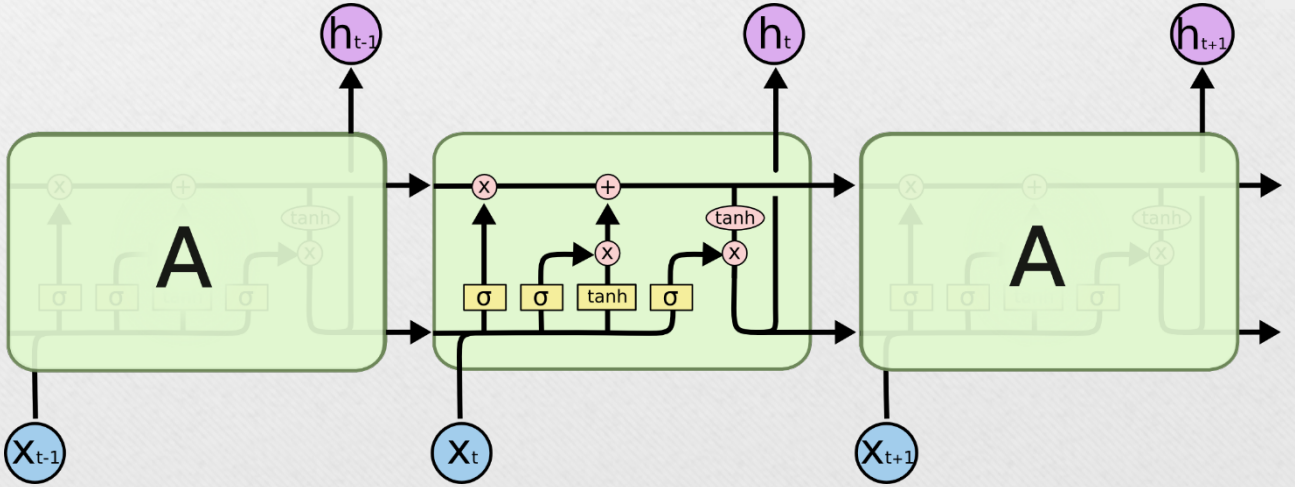


Image Source: colah's blog

LSTM

$$\sigma(x) = \frac{e^x}{1 + e^x}$$

Range: $(0,1)$

$$\tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1}$$

Range: $(-1,1)$

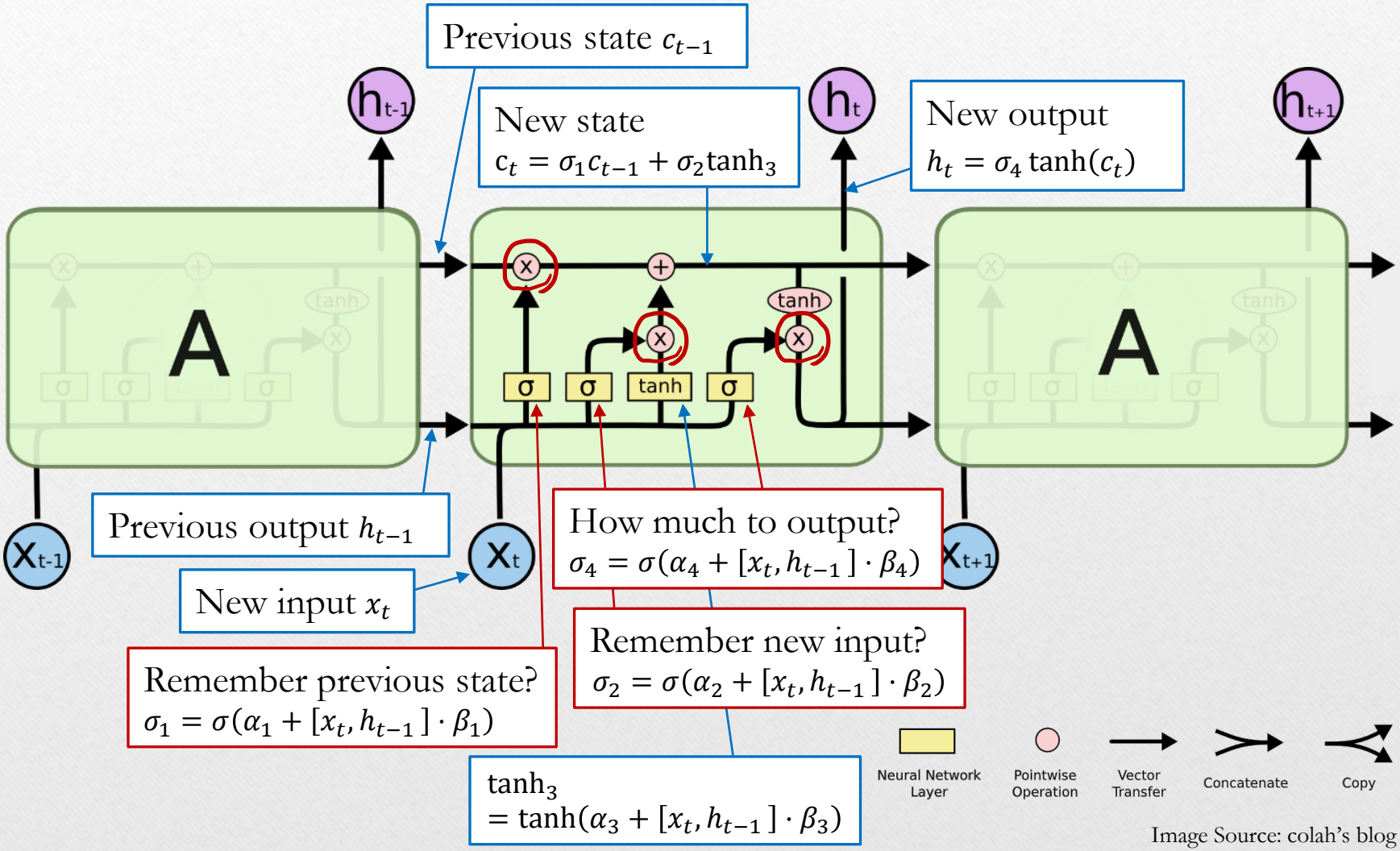


Image Source: colah's blog

Bidirectional RNN

- In a **Bidirectional RNN**, half the neurons in each layer take input in the reverse order.
- Now the network remembers not just the past but also the future. Make sense in NLP, but for time series data?
- In most framework including Keras, the same setting would result in twice the number of neurons in a Bidirectional RNN when compared to a standard RNN.

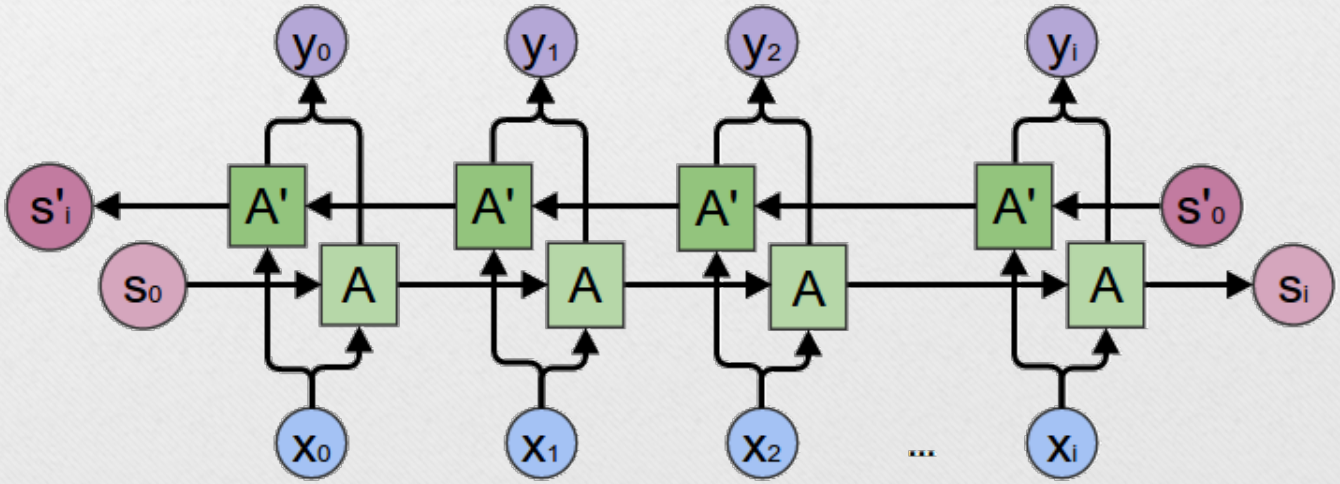


Image Source: colah's blog