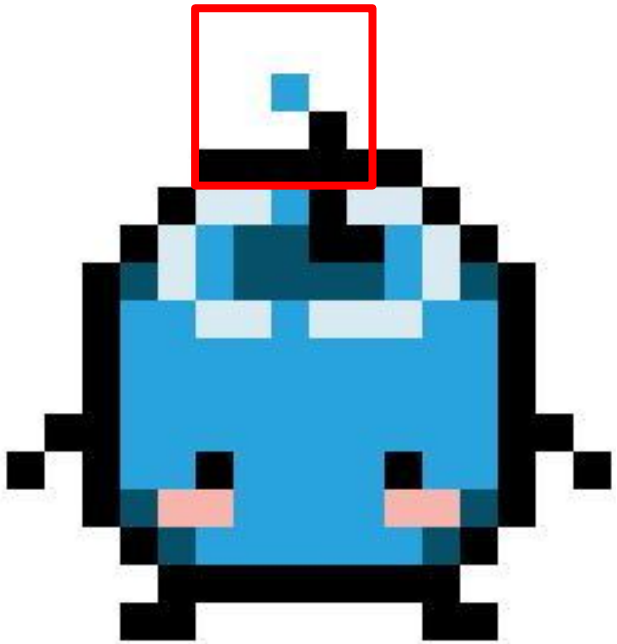
Recurrent Neural Networks I

Dr. Parlett-Pelleriti

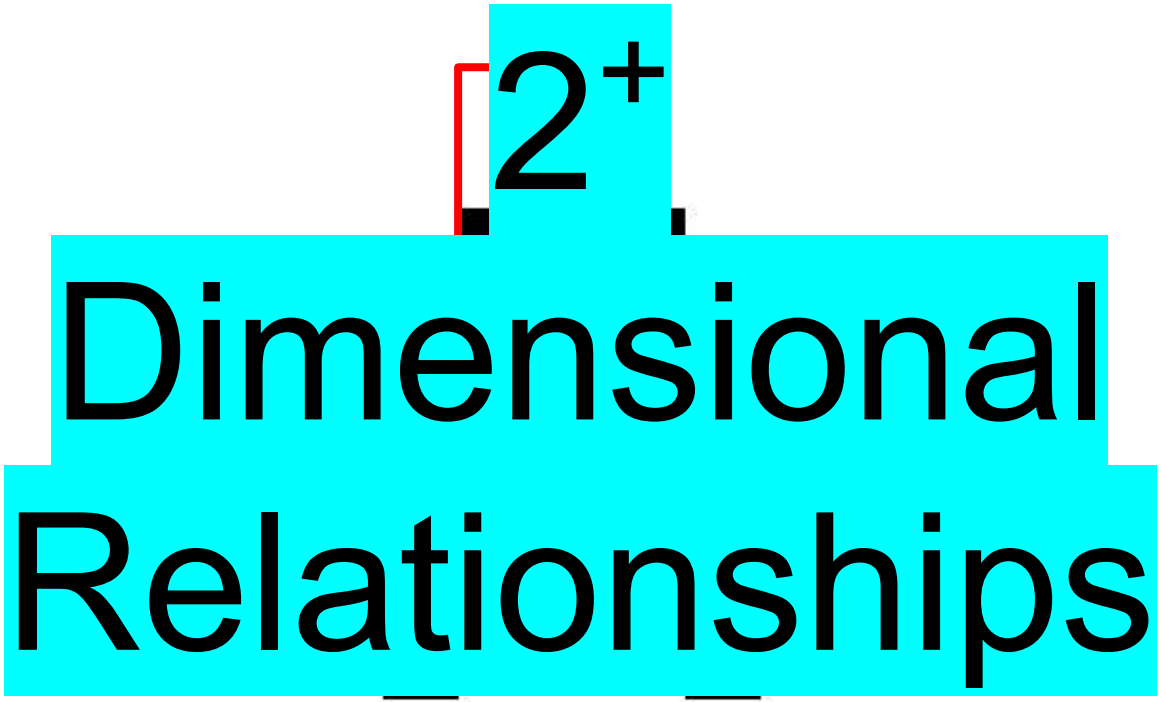
Outline

- Sequential Data
- FF and Conv Architectures
- RNNs
- Vanishing Gradient
- LSTMs
- GRUs

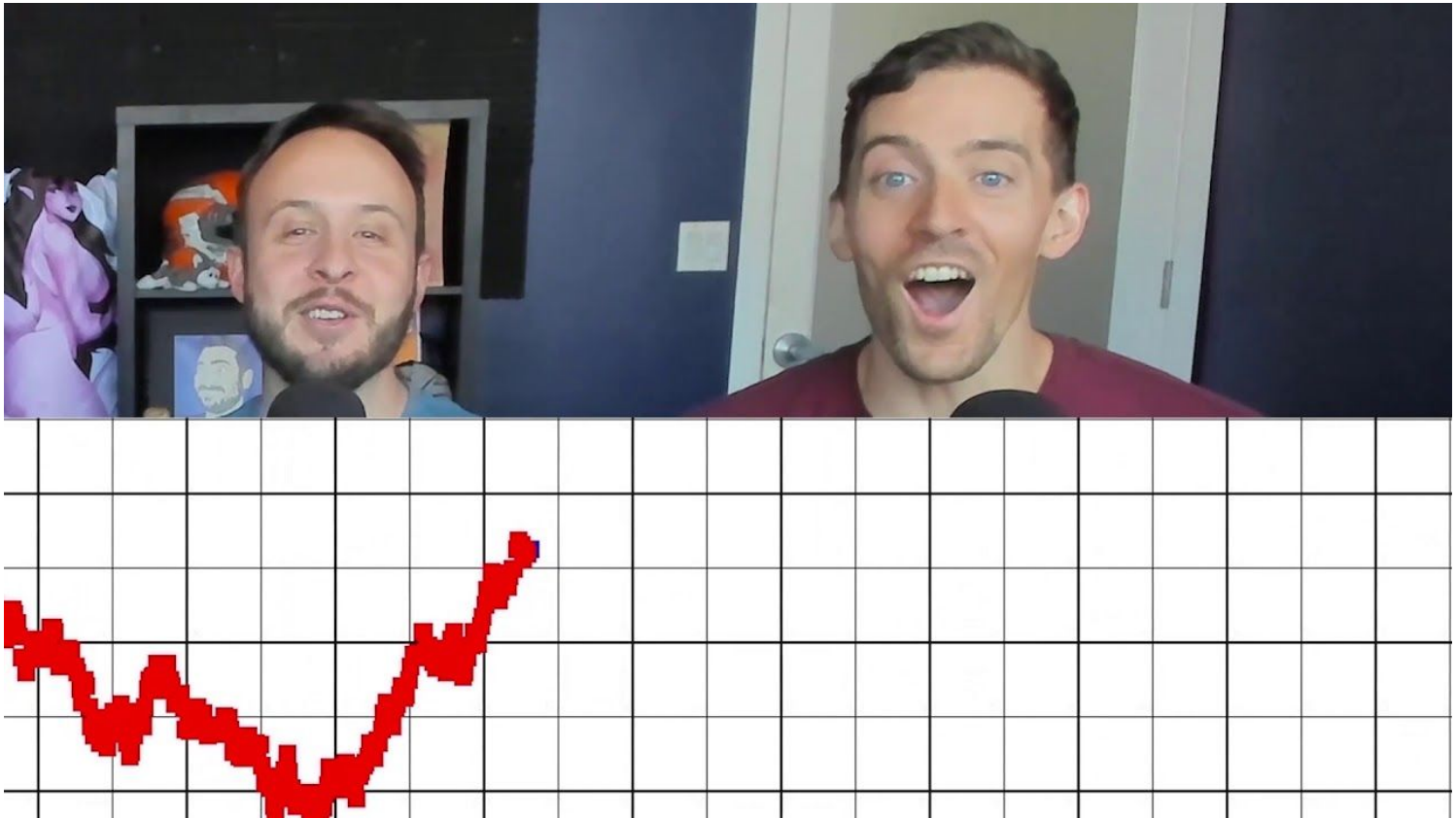
Convolutional Architecture



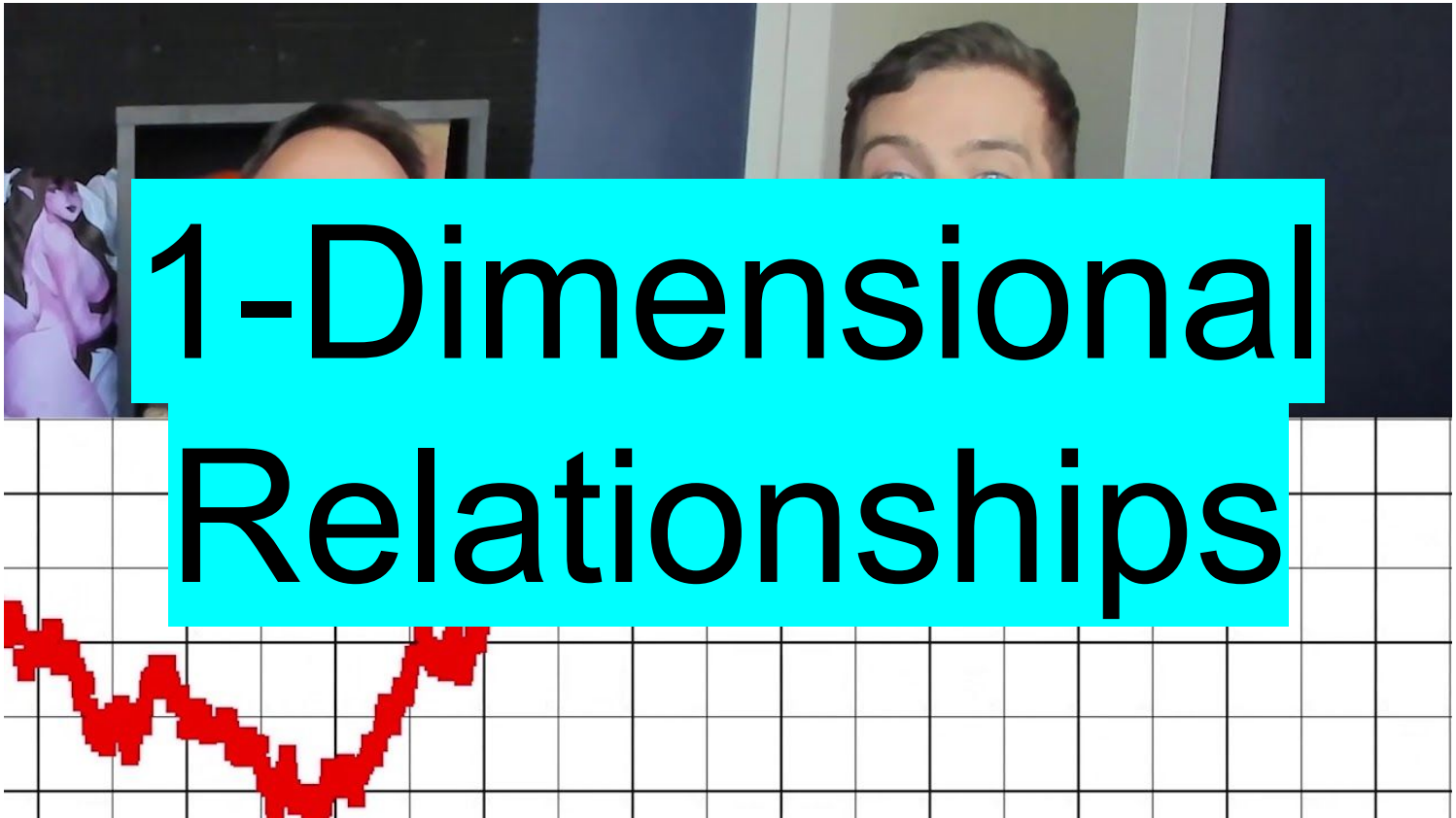
Convolutional Architecture



Sequential Architecture



Sequential Architecture



Sequential Data Examples

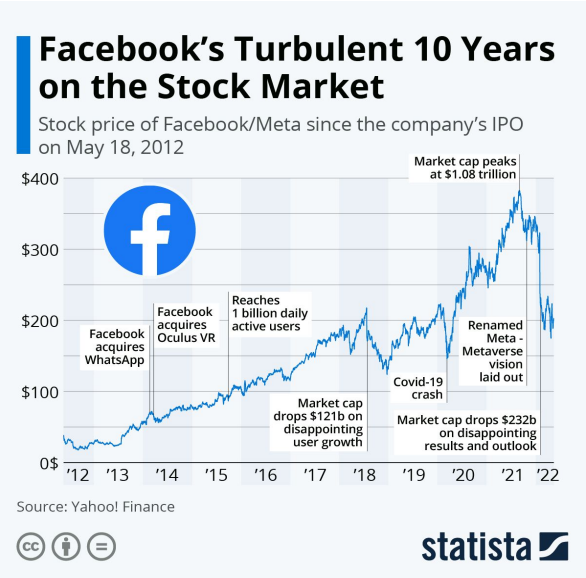


Image from: https://www.statista.com/chart/27473/facebook-stock-price-since-its-ipo/

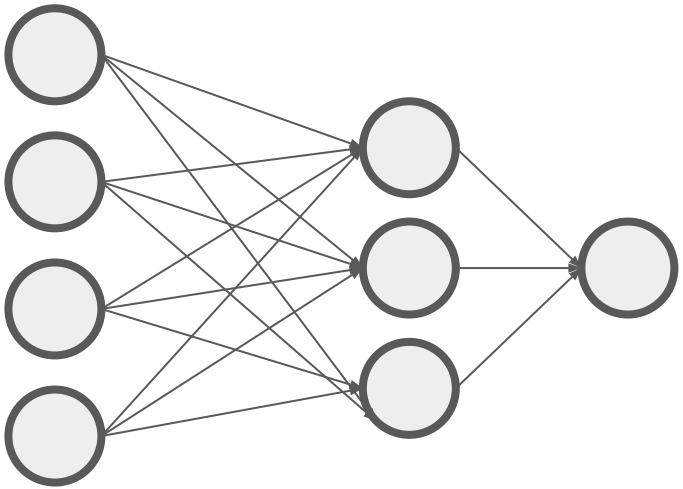
Image from: https://www.theverge.com/23004703/heart-rate-ekg-smartwatch

Image from: https://pxhere.com/en/photo/890274

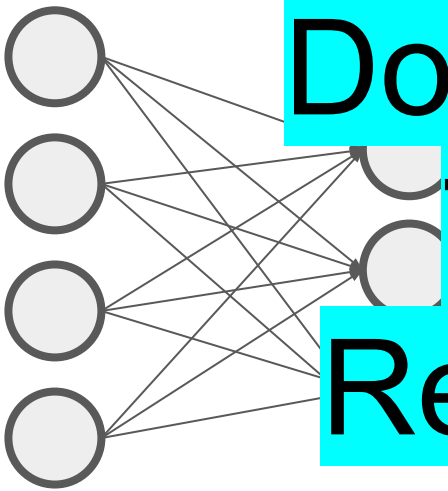
Things To Do With Sequential Data

- Forecasting
- Sequence Classification
- Machine Translation
- Anomaly Detection
- Summarization
- Sequence Generation

Feed Forward Model

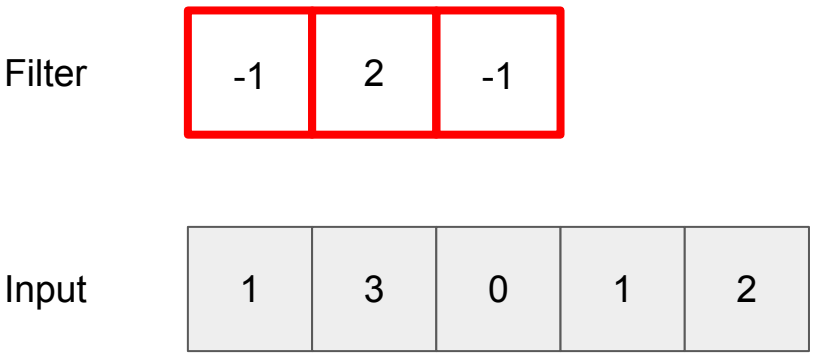


Feed Forward Model

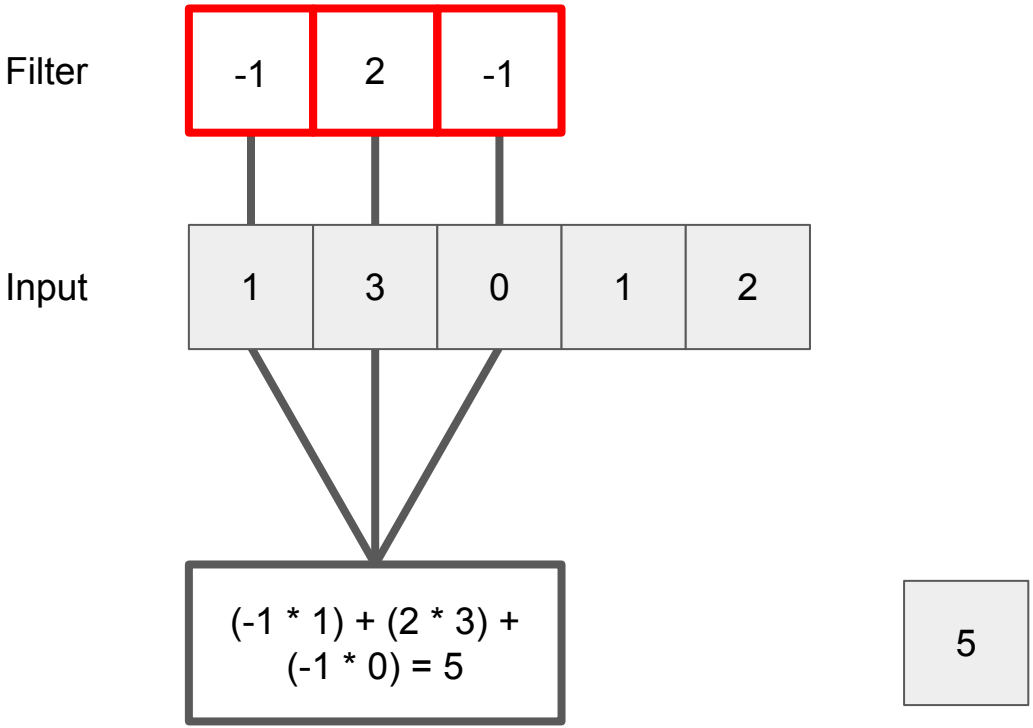


Doesn’t Consider Time Based Relationships of Inputs

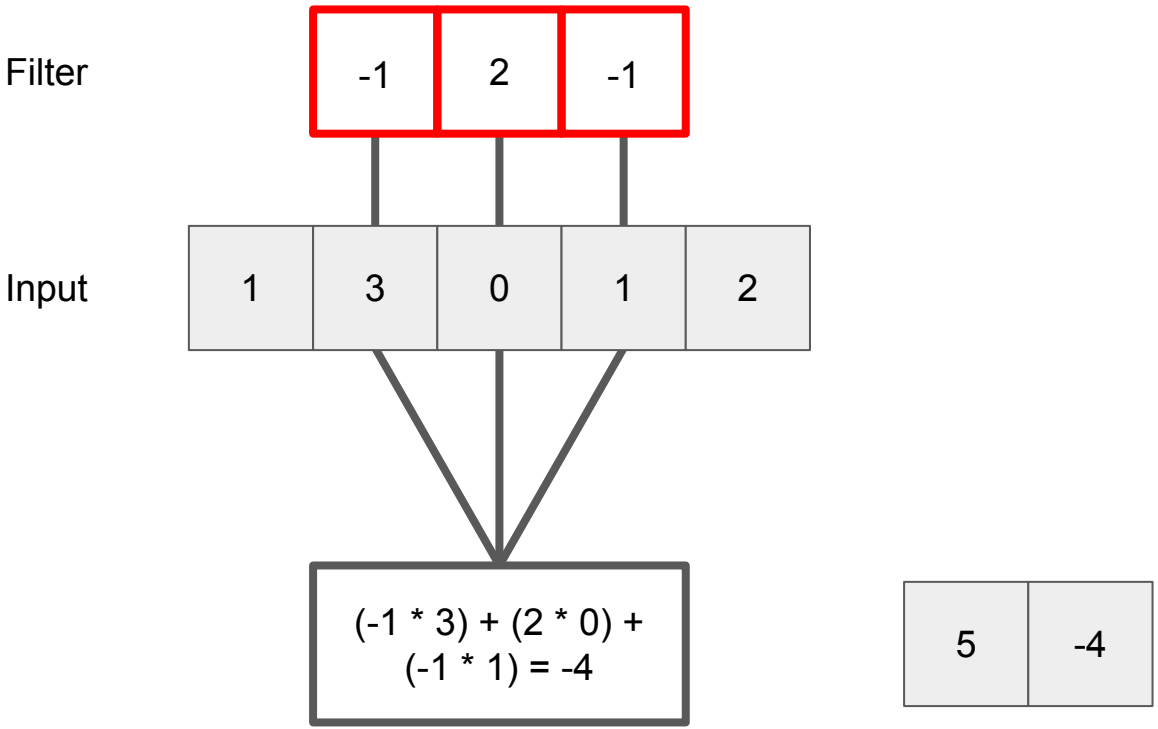
1-D Convolutional Layers



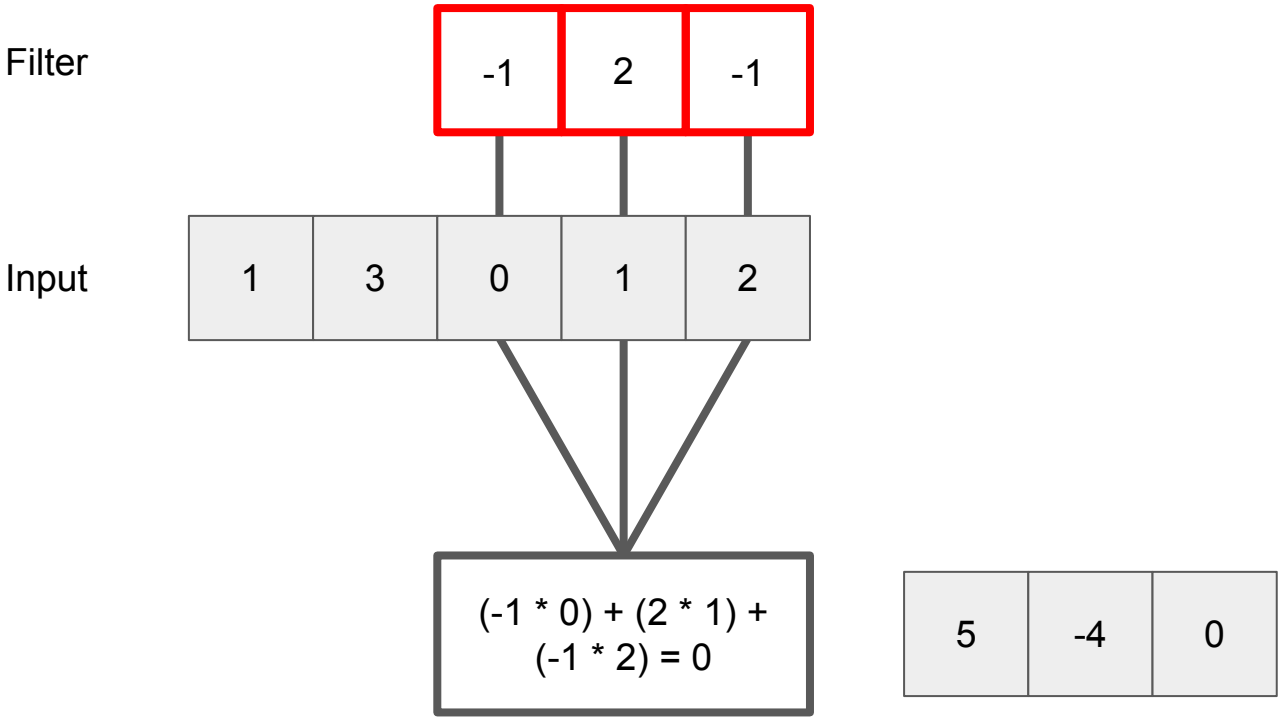
1-D Convolutional Layers



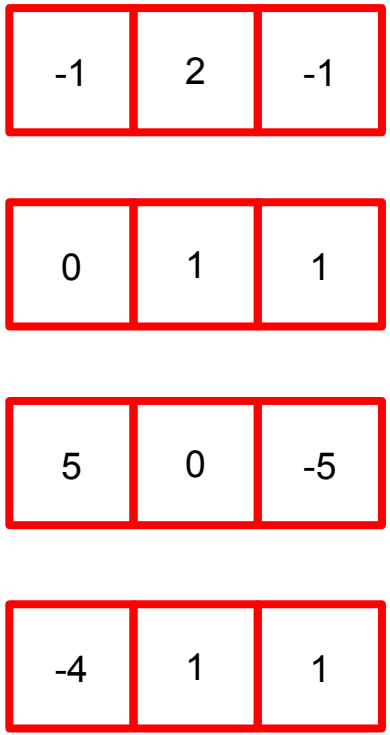
1-D Convolutional Layers



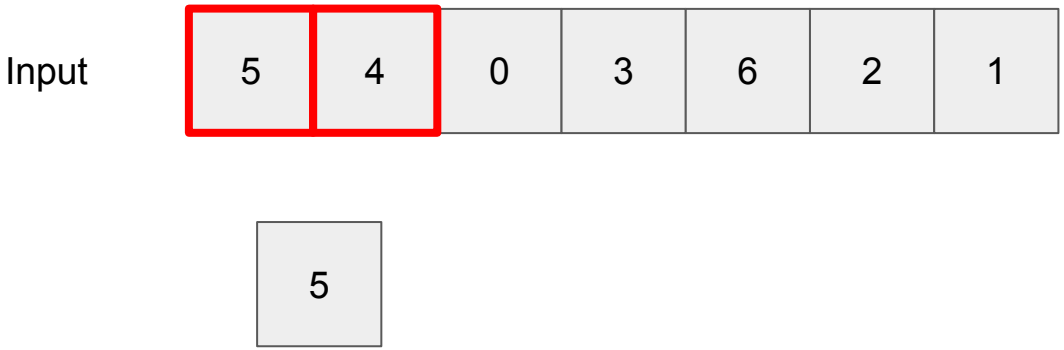
1-D Convolutional Layers



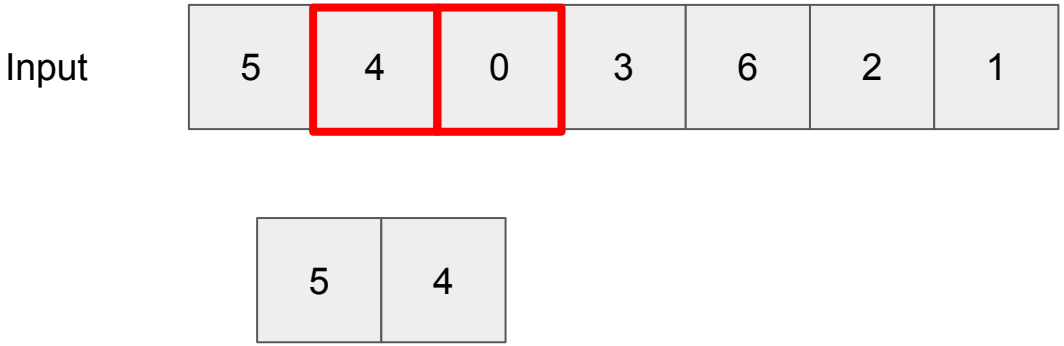
1-D Convolutional Layers



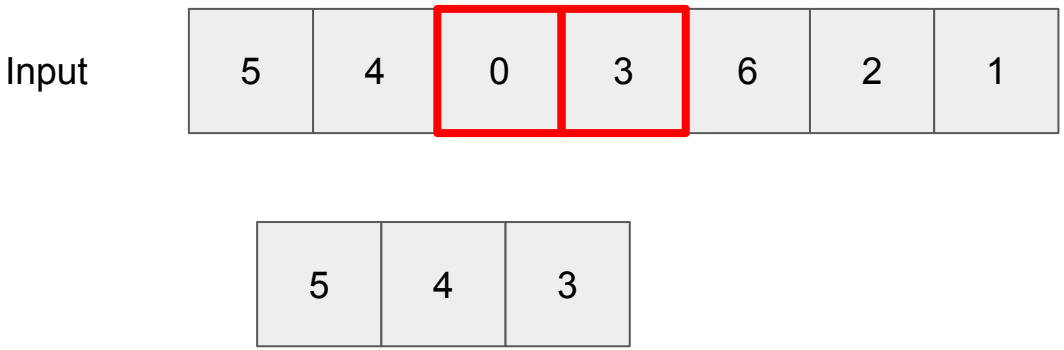
1-D Max Pooling Layers



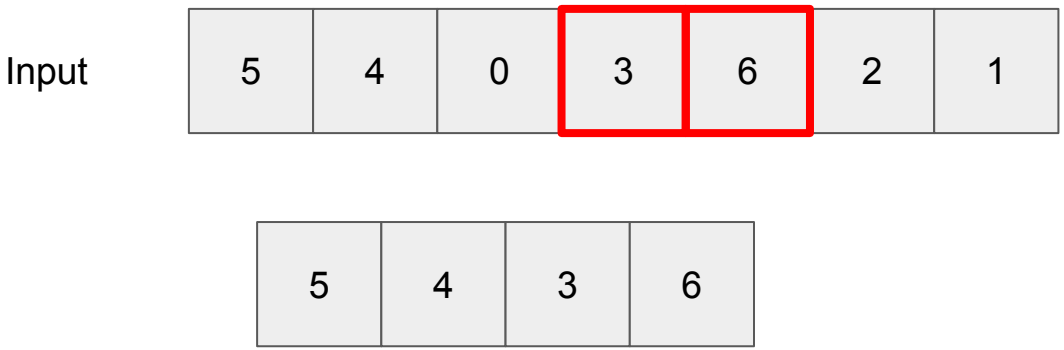
1-D Max Pooling Layers



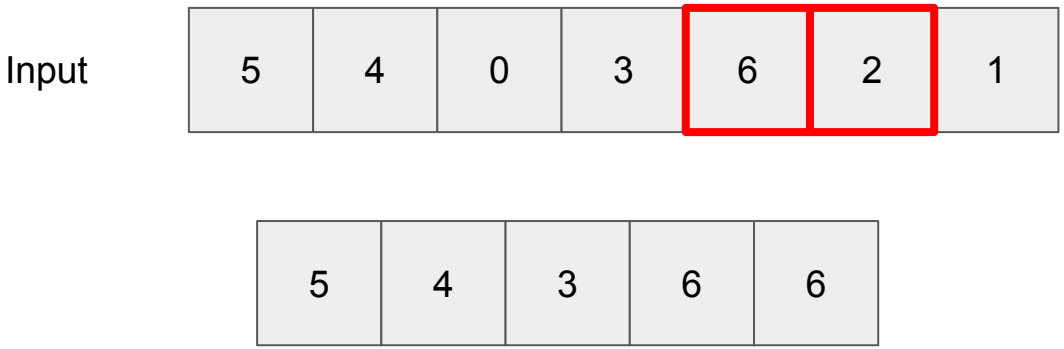
1-D Pooling Layers



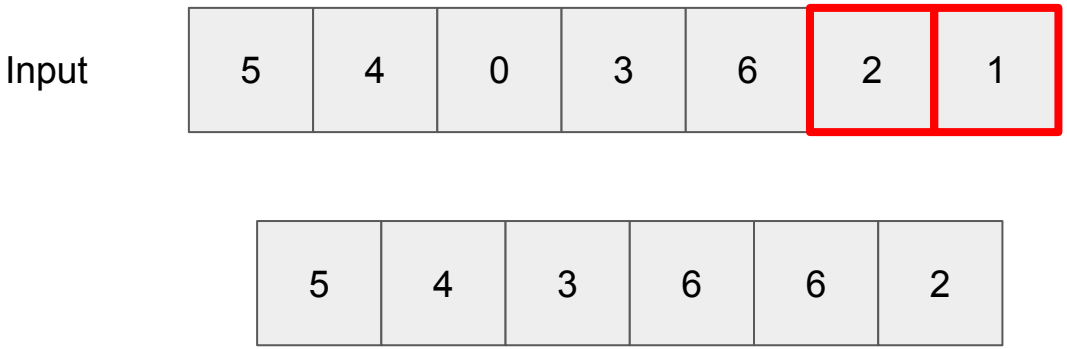
1-D Max Pooling Layers



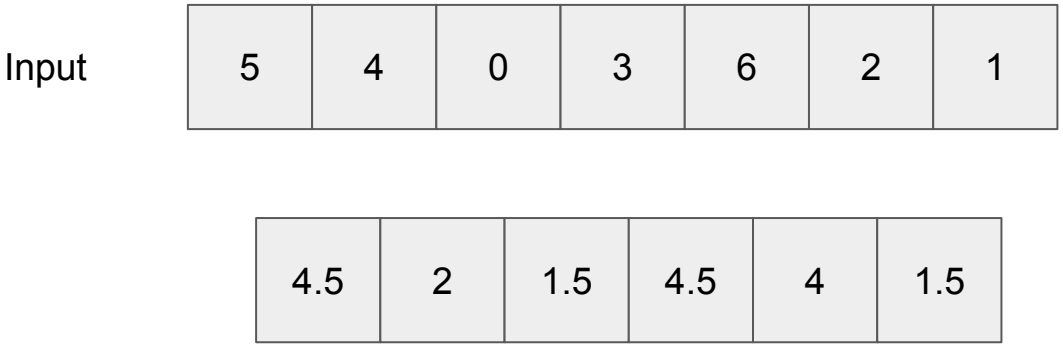
1-D Max Pooling Layers



1-D Max Pooling Layers



1-D Average Pooling Layers

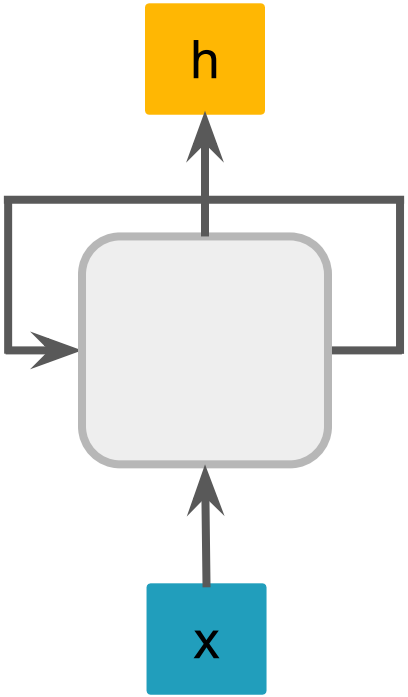


1-D Conv and Pooling

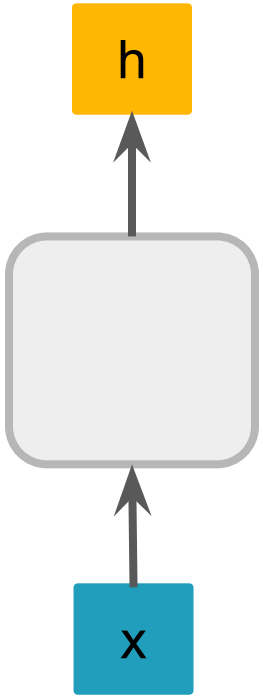
- Like with spatial data, filters assume translational invariance
- Pooling can interfere with sequential information

Recurrent Architectures

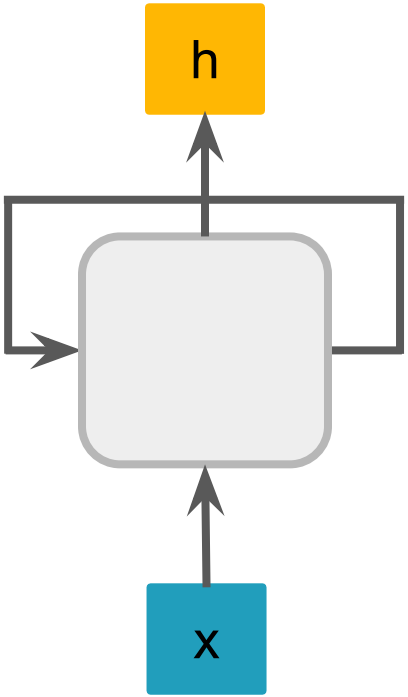
Recurrent Structure Generally



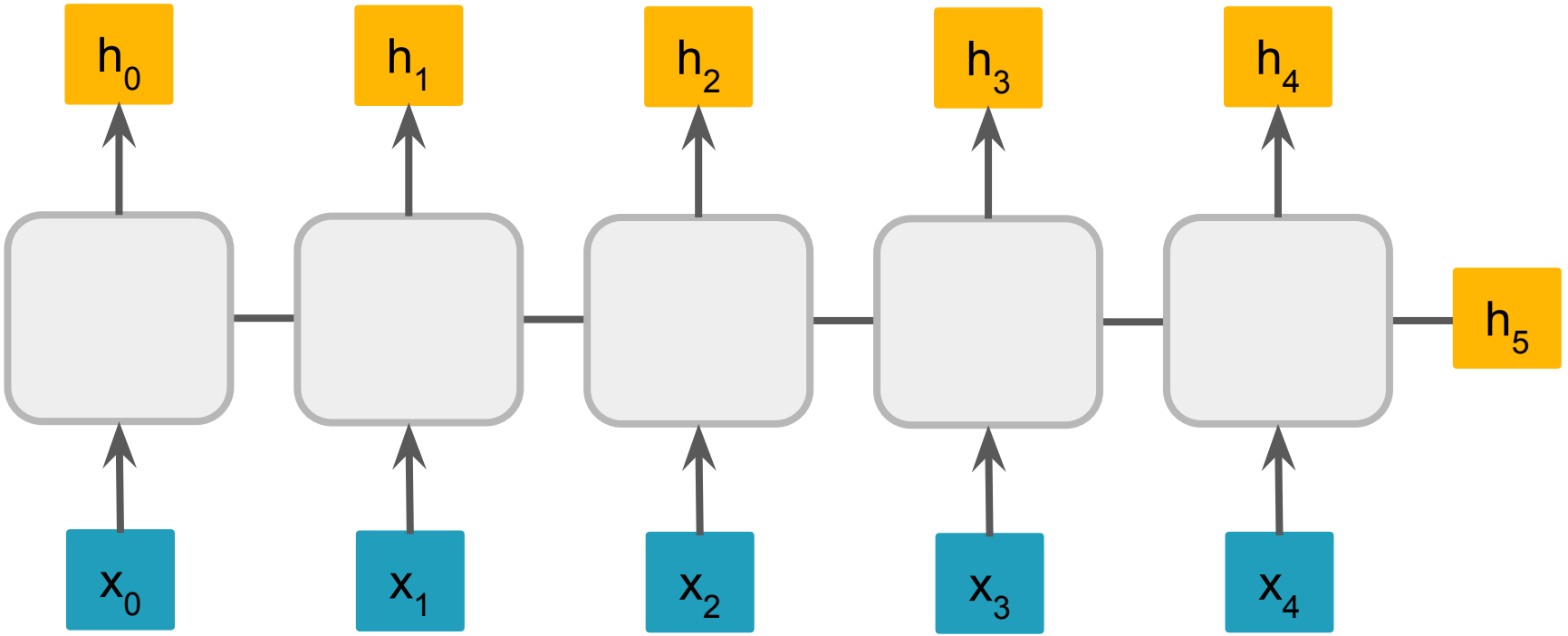
Recurrent Structure Generally



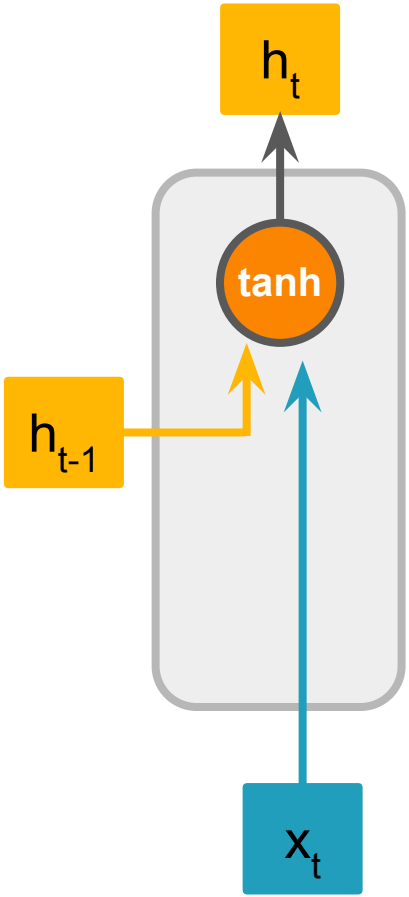
Recurrent Structure Generally



Recurrent Structure Unrolled

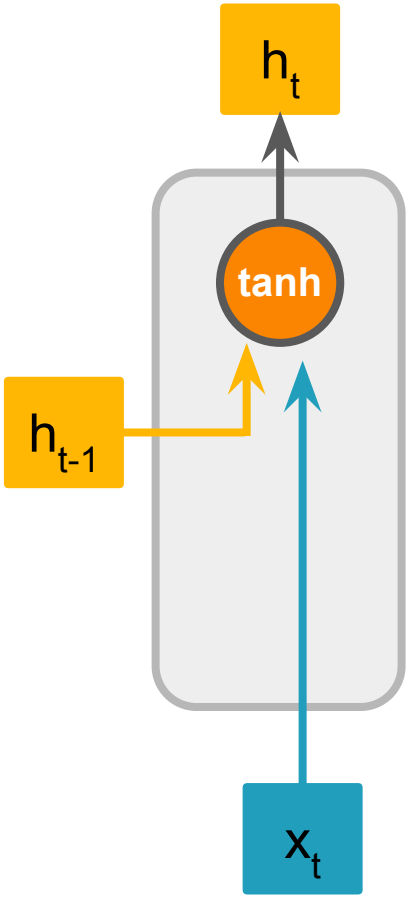


Simple RNN



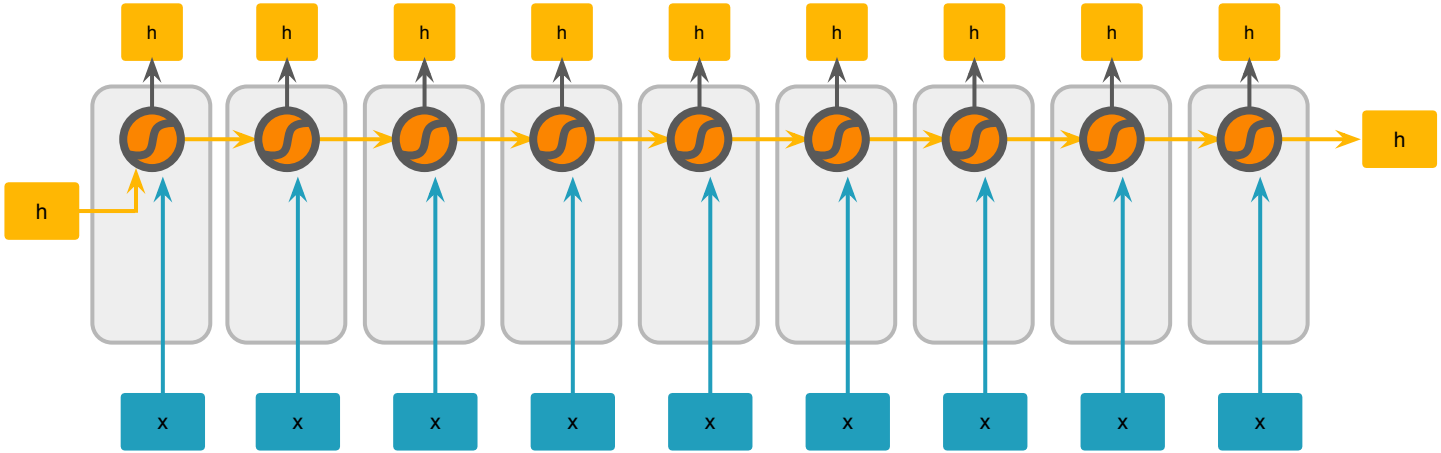
$$h_t = \tanh(W \cdot x_t + U \cdot h_{t-1} + b)$$

Simple RNN



$$h_t = \tanh(W \cdot [h_{t-1}, x_t] + b)$$

Simple RNN Unrolled



Bidirectional RNN

The quick brown _____ jumped over the lazy dog

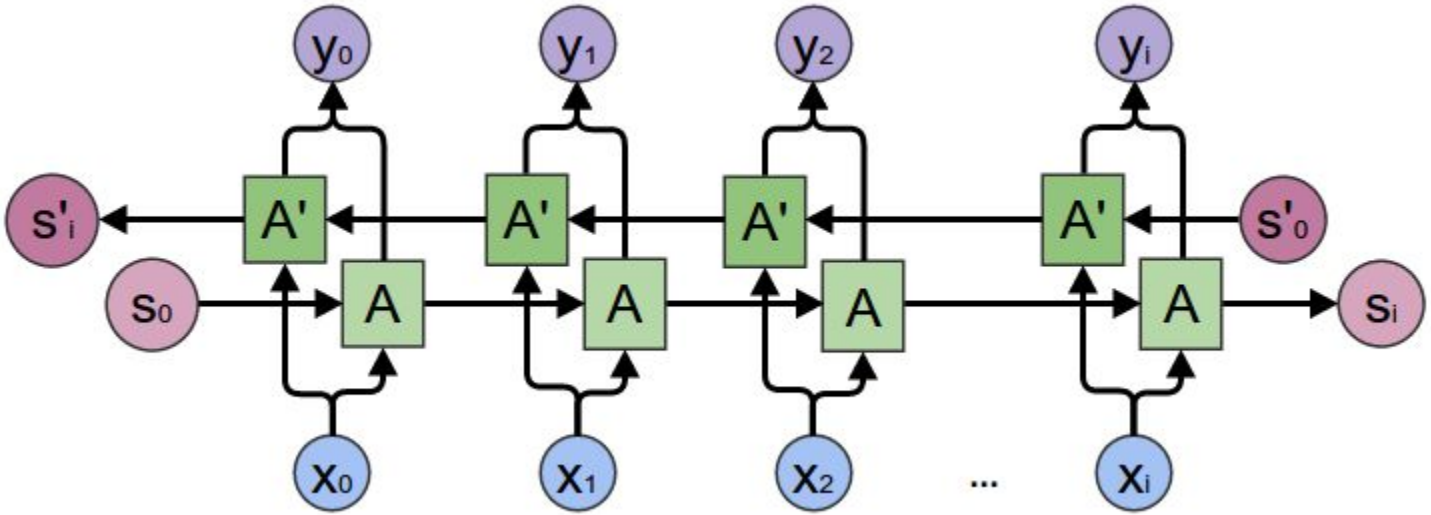


Image from: https://colah.github.io/posts/2015-09-NN-Types-FP/

Bidirectional RNN

The quick brown _____ jumped over the lazy dog

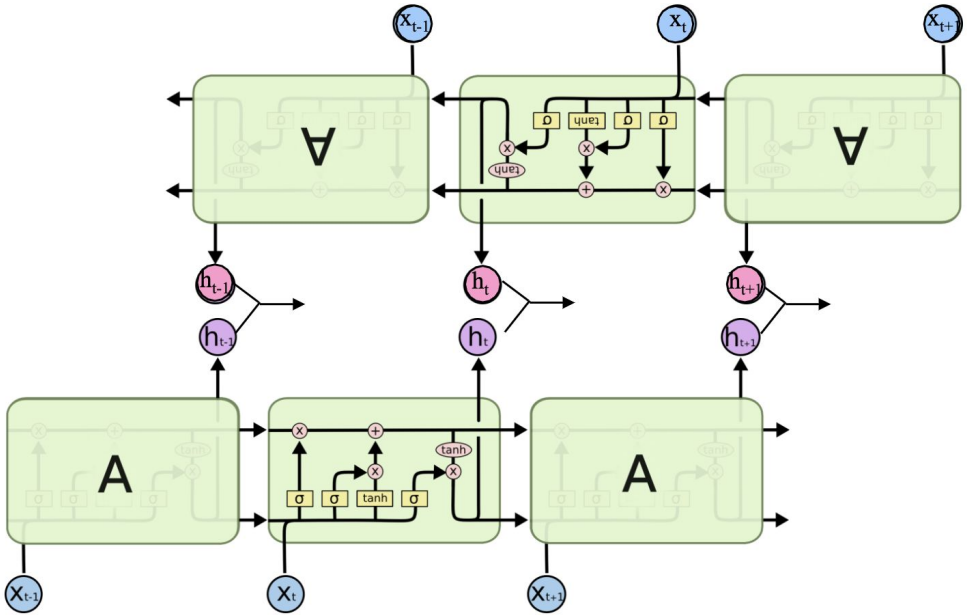


Image from: http://colah.github.io/posts/2015-08-Understanding-LSTMs/ and **CS 388: Natural Language Processing: LSTM Recurrent Neural Networks by Raymond J Mooney**