Module 11
Overview of Deep Learning

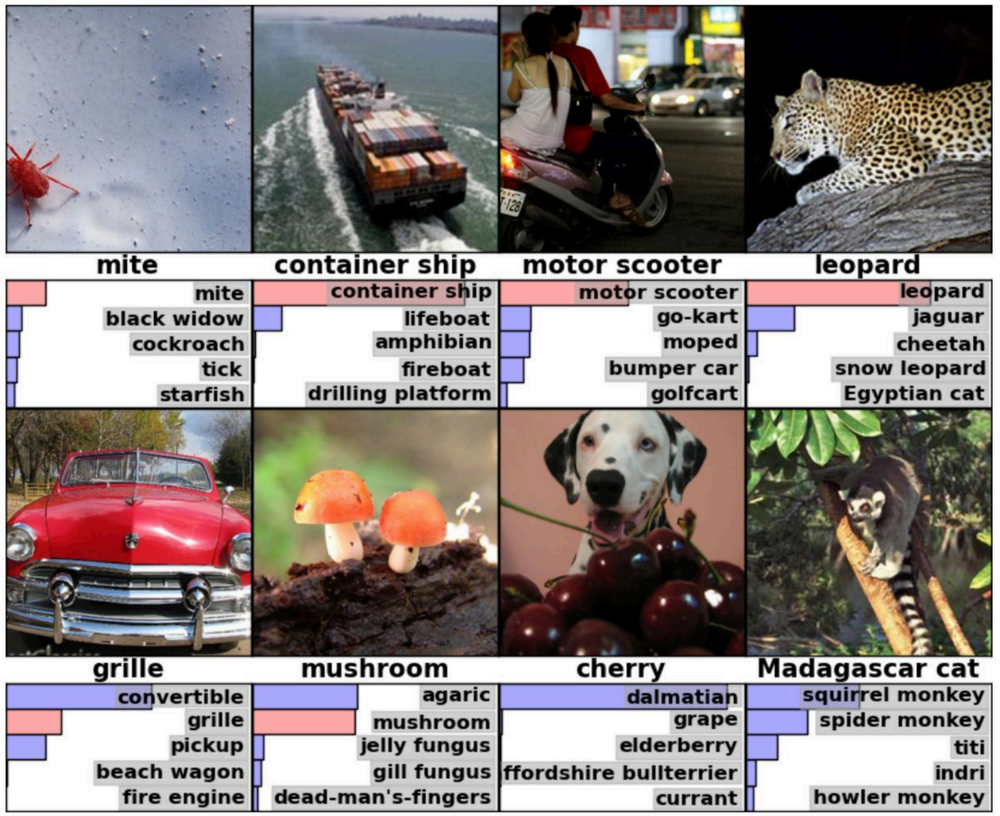
Why?

Why?

ImageNet Challenge

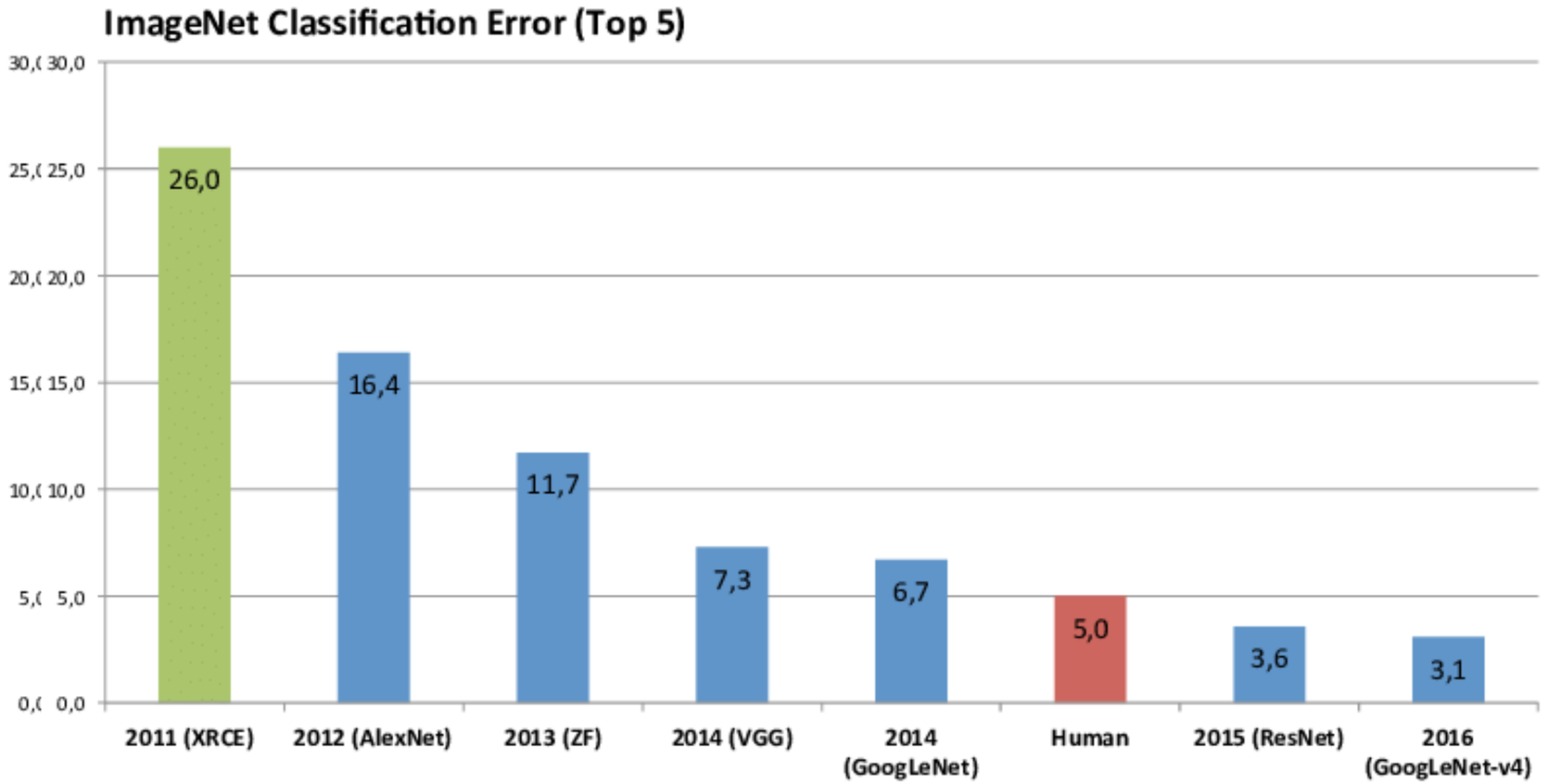
IMAGENET

- 1,000 object classes (categories).
- Images:
  - 1.2 M train
  - 100k test.



(Source: Xavier Giro-o-Nieto)

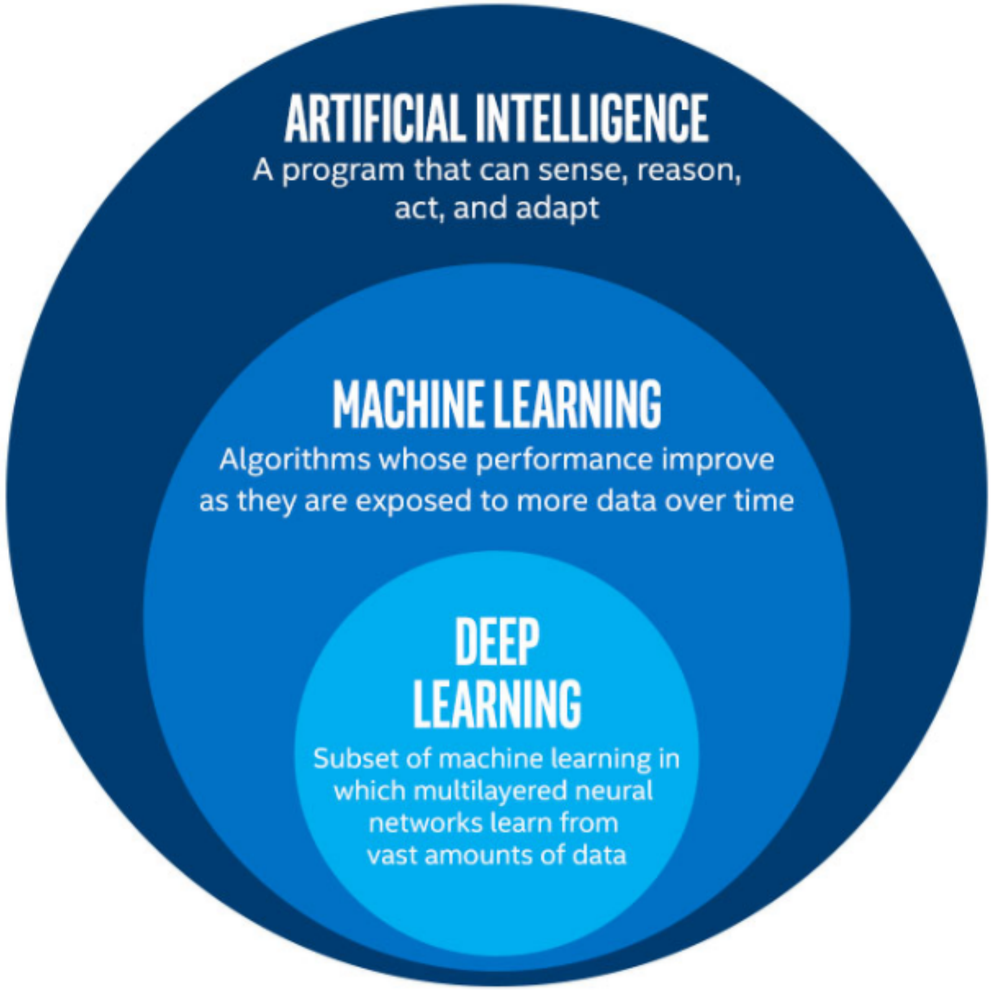
Why?



https://www.researchgate.net/figure/Winner-results-of-the-ImageNet-large-scale-visual-recognition-challenge-LSVRC-of-the_fig7_324476862

What is Deep Learning?

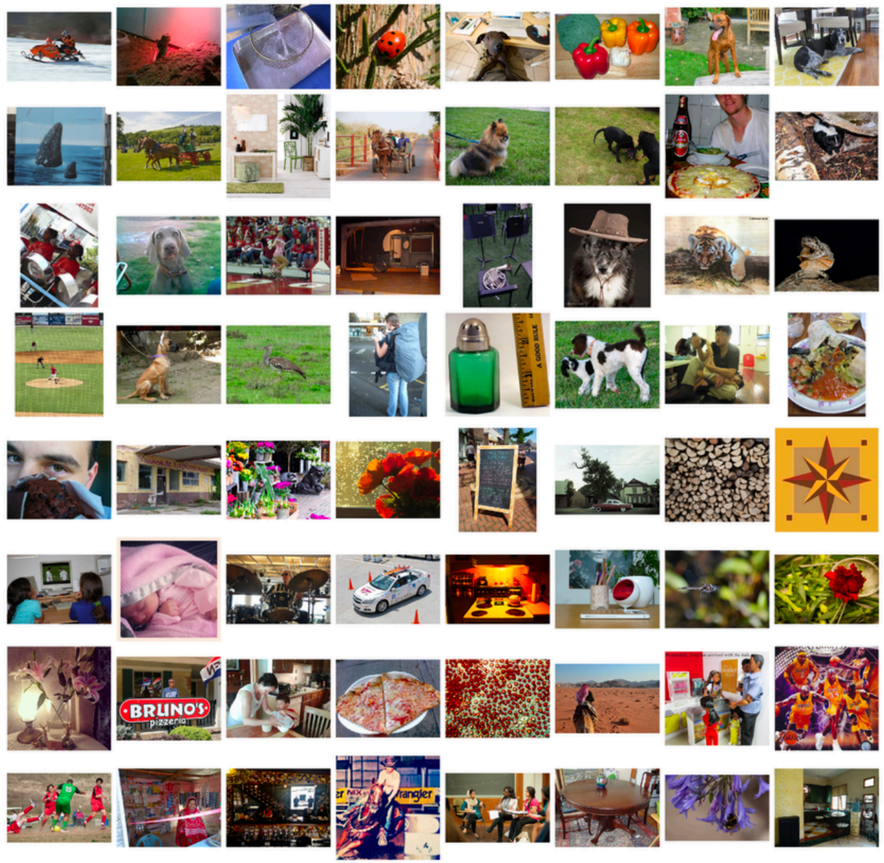
What is Deep Learning?



https://towardsdatascience.com/cousins-of-artificial-intelligence-dda4edc27b55

What is Deep Learning?

ImageNet



Traditional Computer Vision

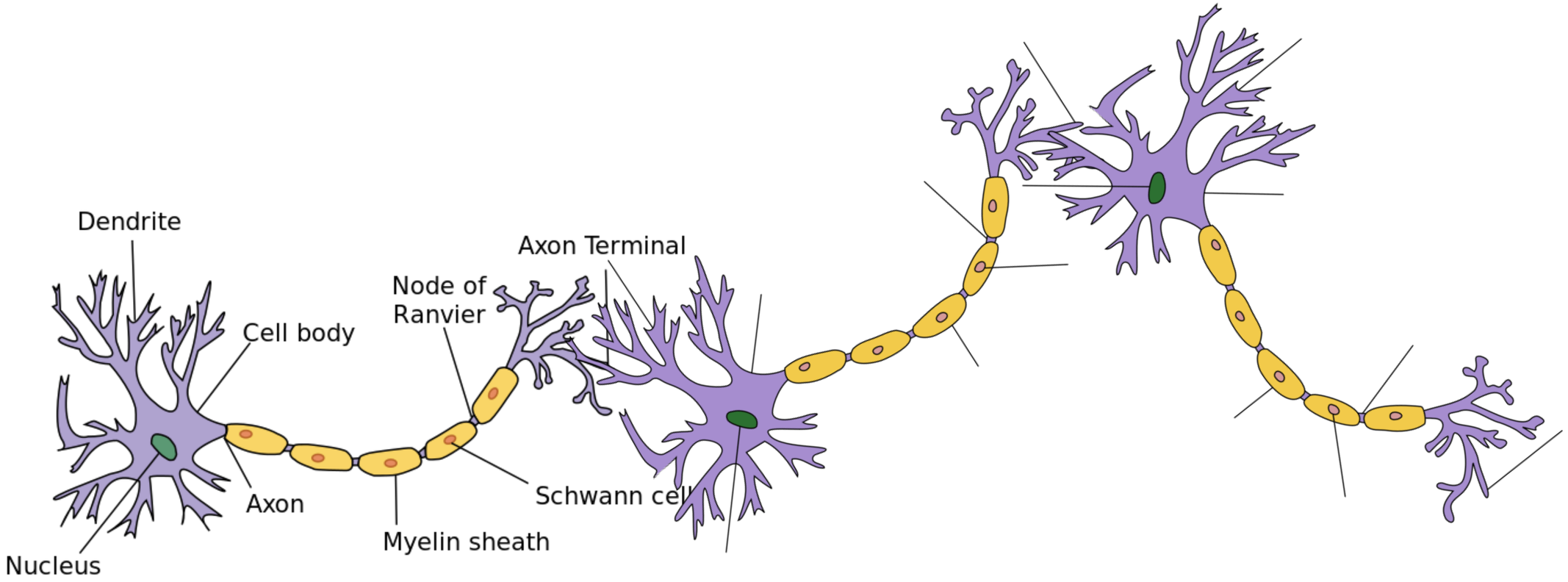
Hand-crafted Feature Extractors → Trainable Classifier

Deep Learning

Low-Level Features → Mid-Level Features → High-Level Features → Trainable Classifier

Neural Networks

Neural Networks



https://en.wikipedia.org/wiki/Neuron

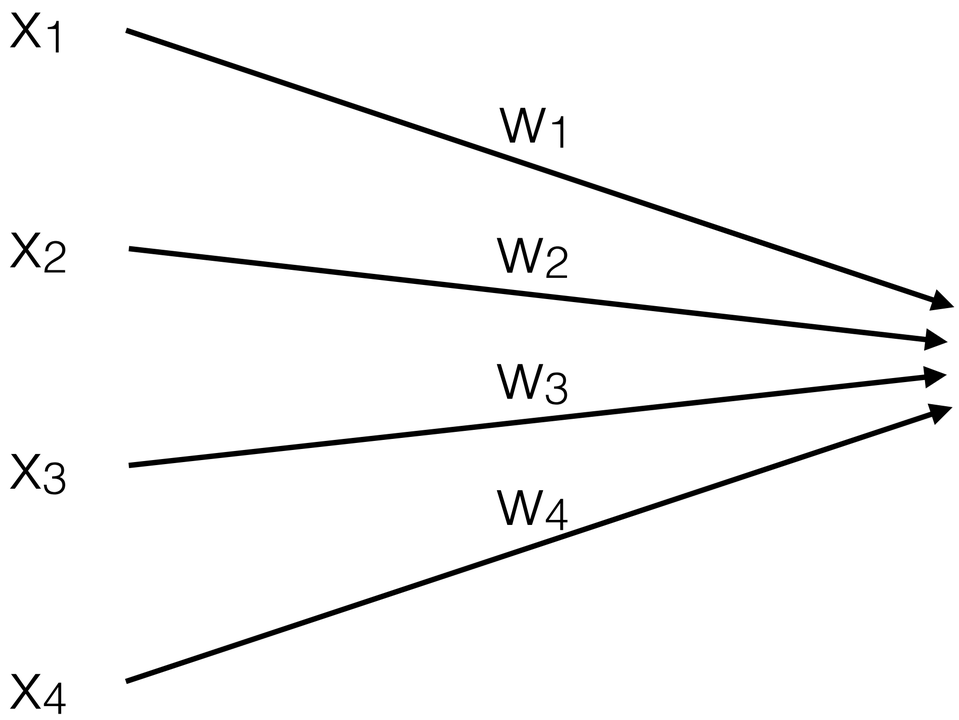
Neural Networks

Inputs

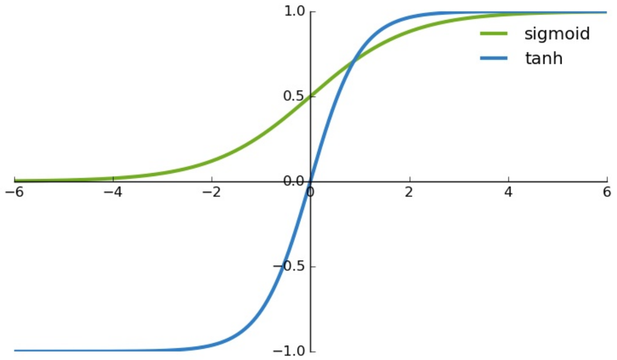
Weights

Summation

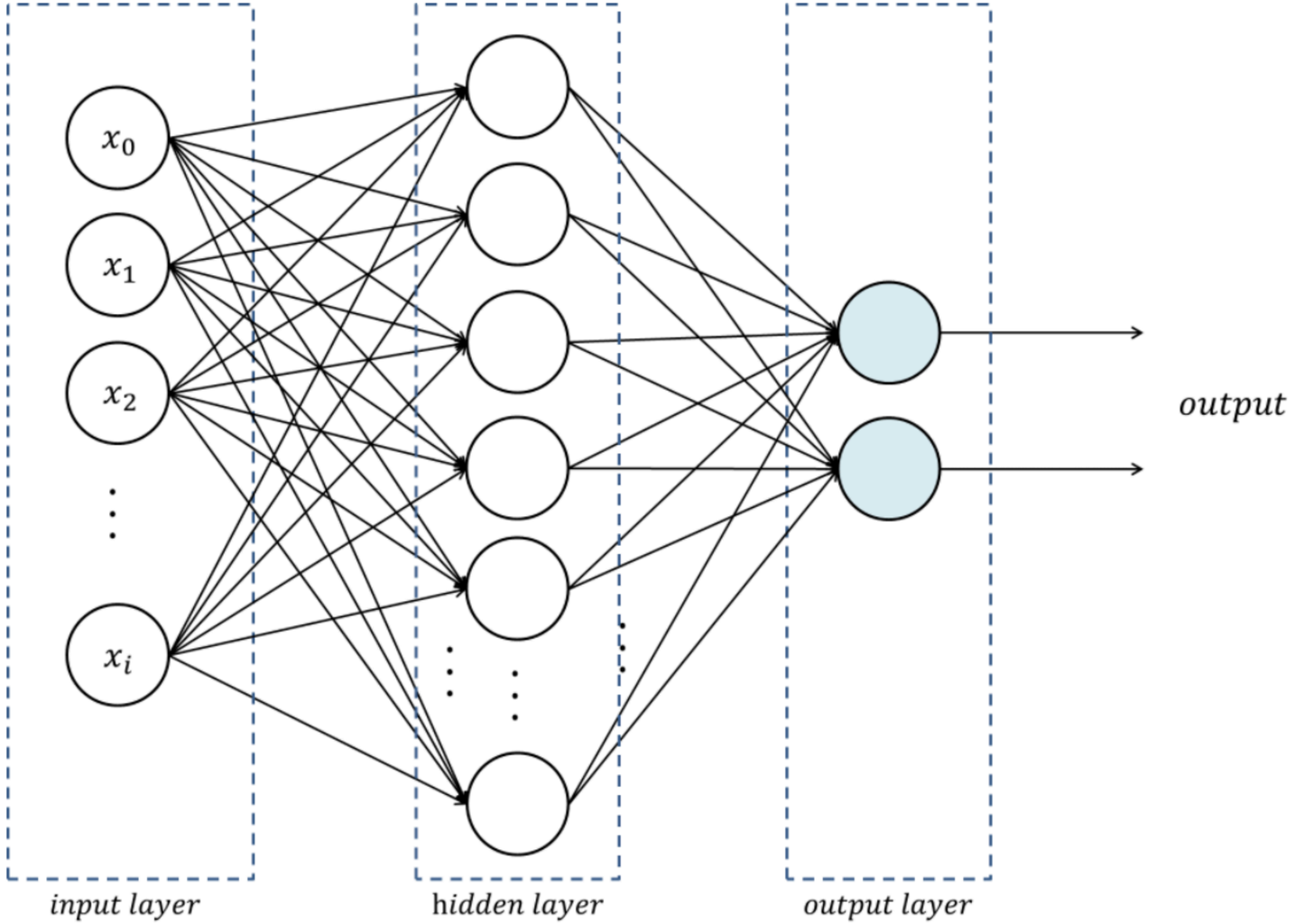
Activation Function



$$\sum x_i w_i \longrightarrow f(x)$$



Multi-layer Perceptron



https://www.cc.gatech.edu/~san37/post/dlhc-fnn/

Neural Network Demo

https://playground.tensorflow.org/

Deep Neural Networks

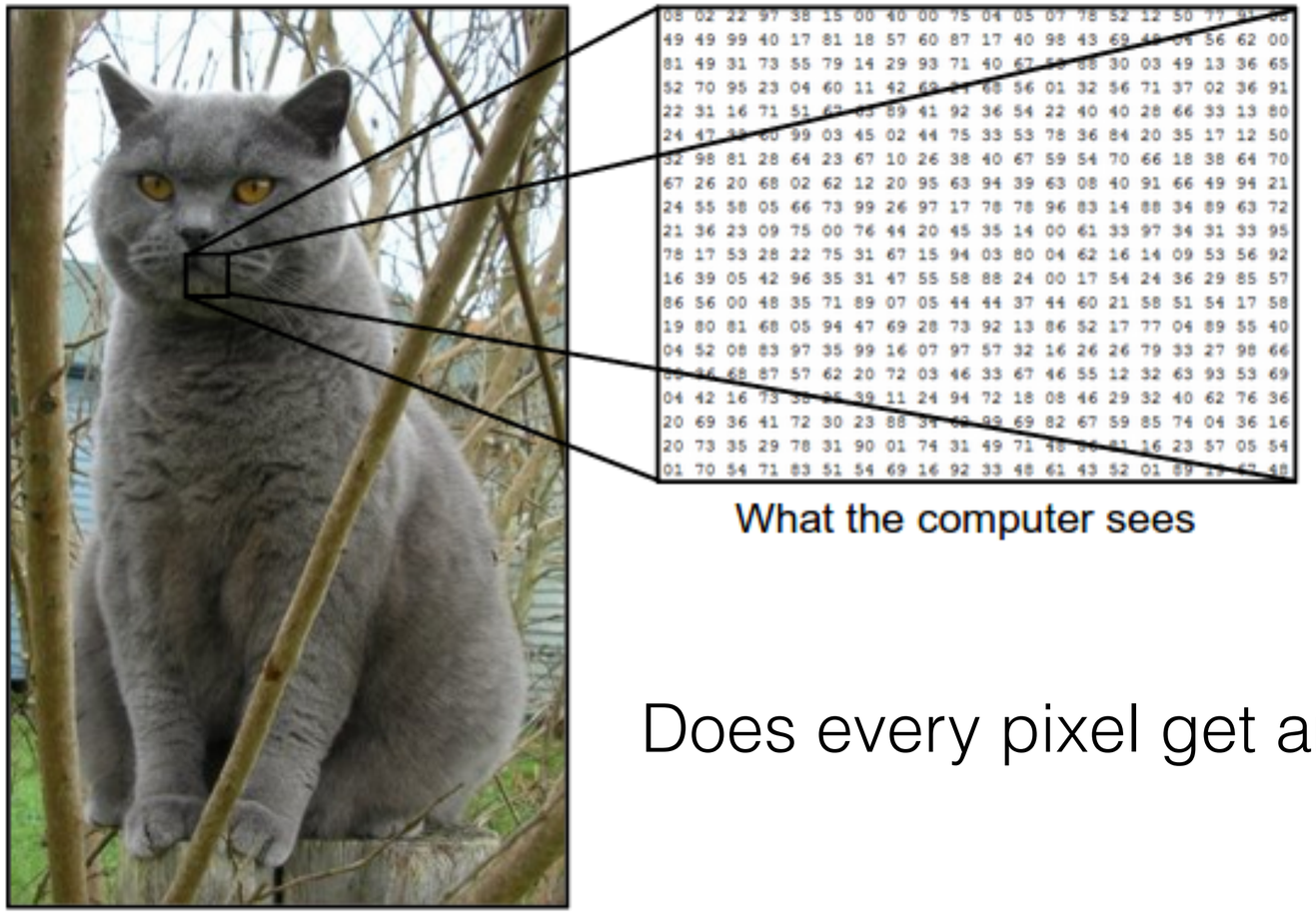
Two types

- Convolutional
- Recurrent

Convolutional Neural Networks

The *Real* Deep Neural Networks

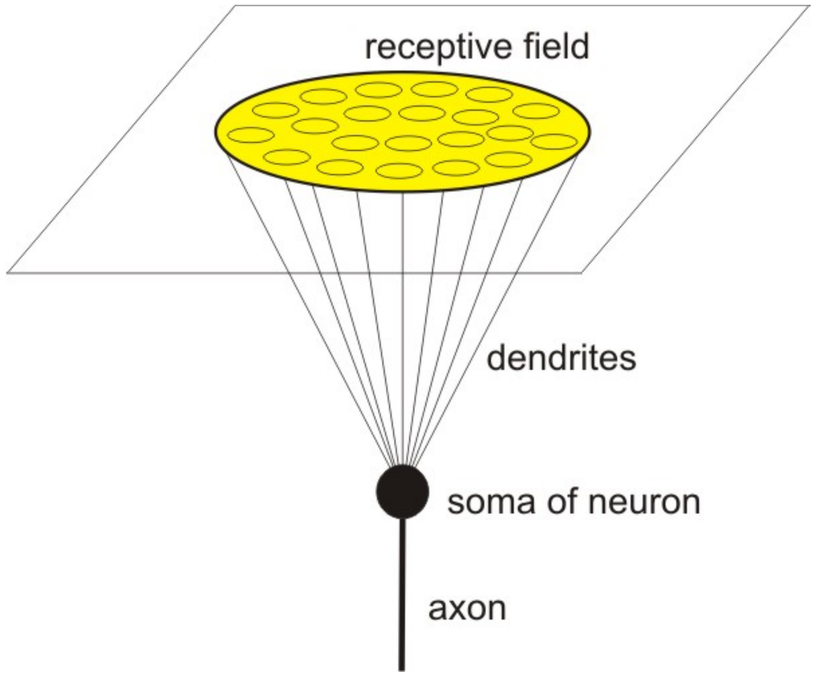
Convolution

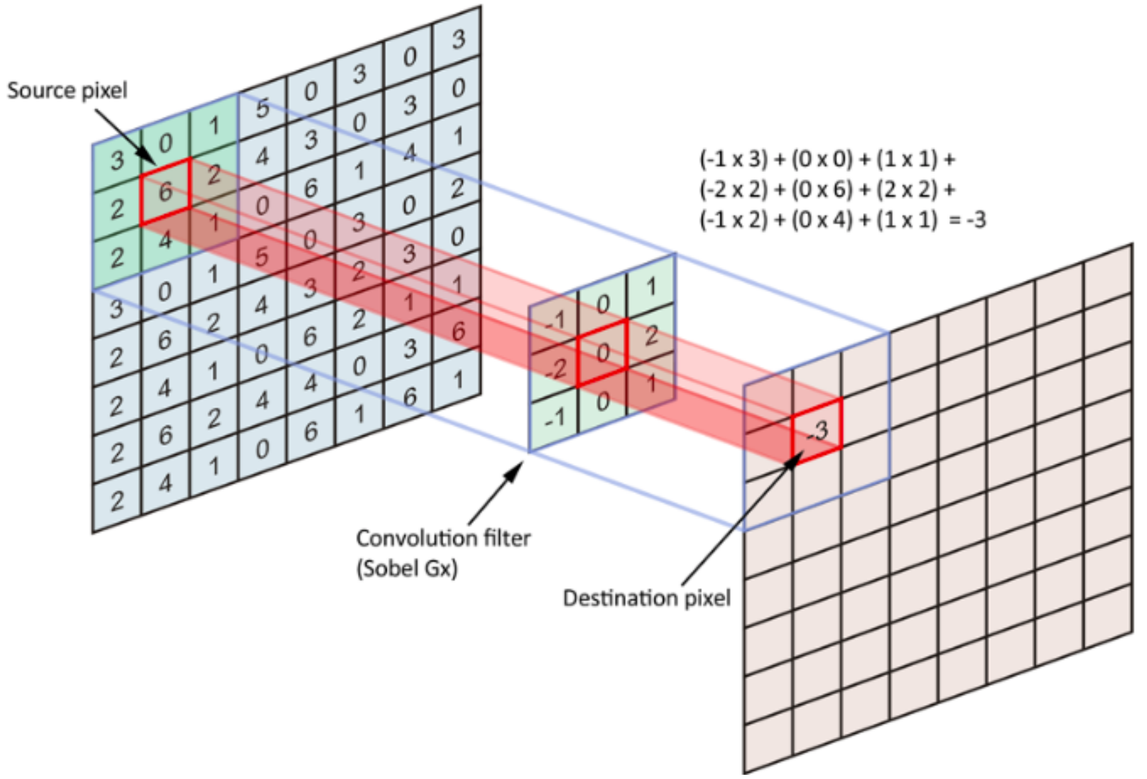


Does every pixel get an input neuron?

http://cs231n.github.io/classification/

Convolution



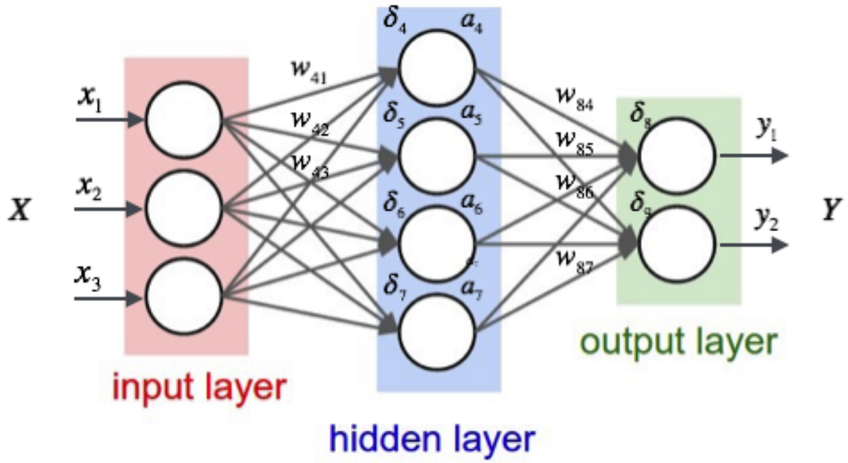


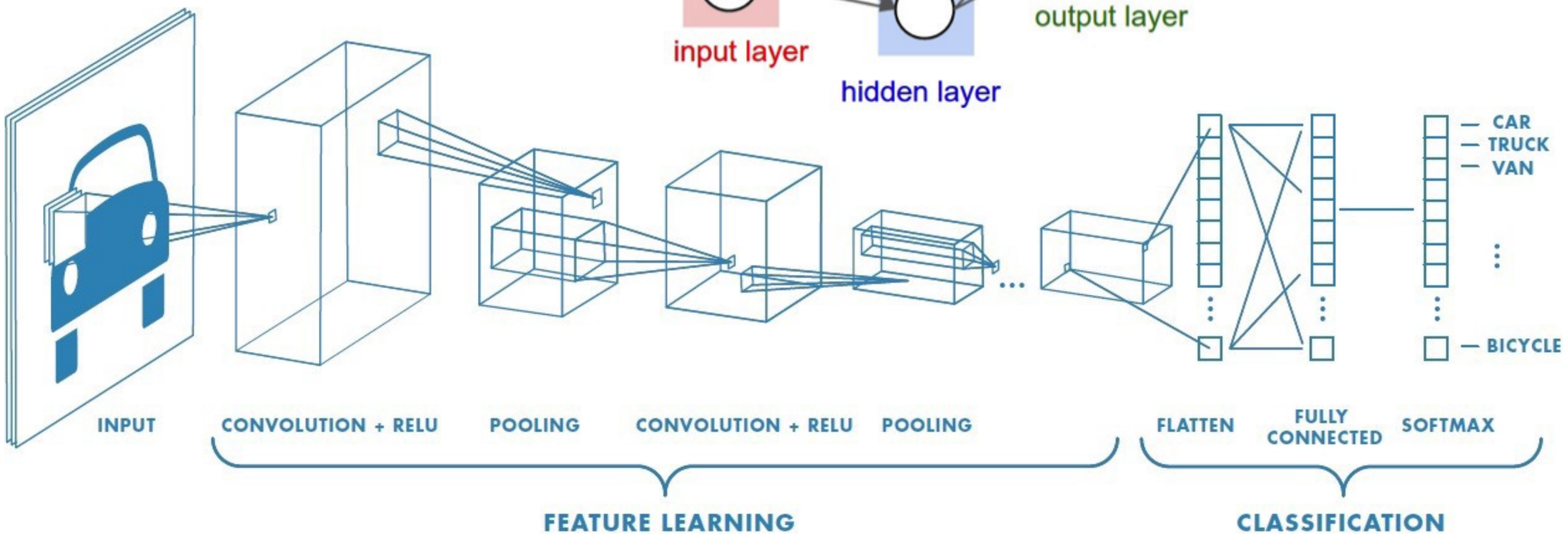
http://neuroclusterbrain.com/neuron_model.html

https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2

gtkcyber.com

Convolution

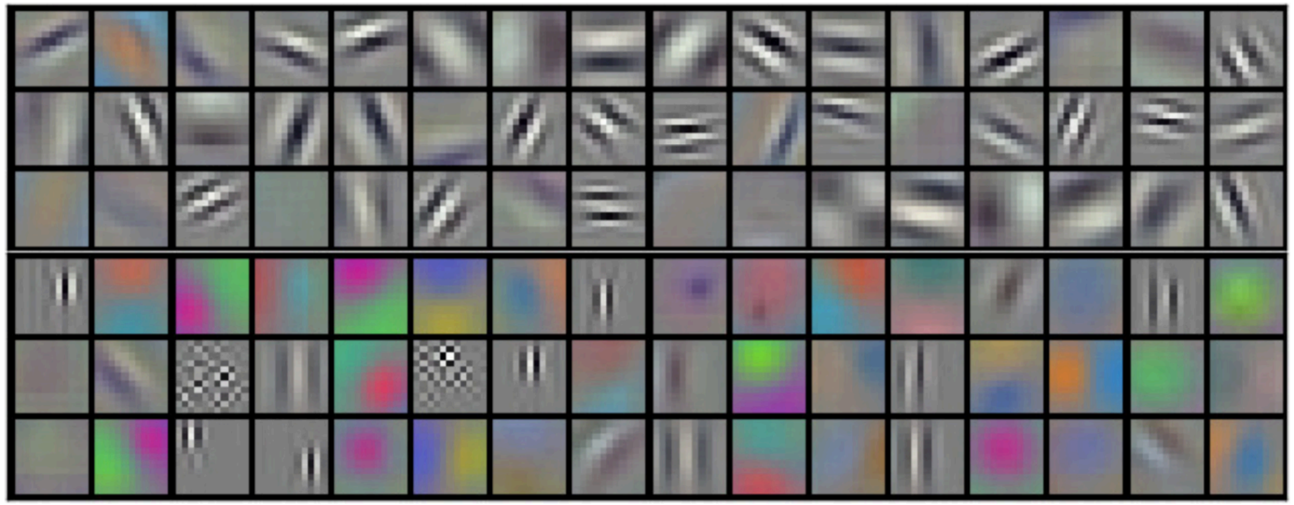




https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

https://medium.com/@curiousily/tensorflow-for-hackers-part-iv-neural-network-from-scratch-1a4f504dfa8

Convolution - Learned Filters



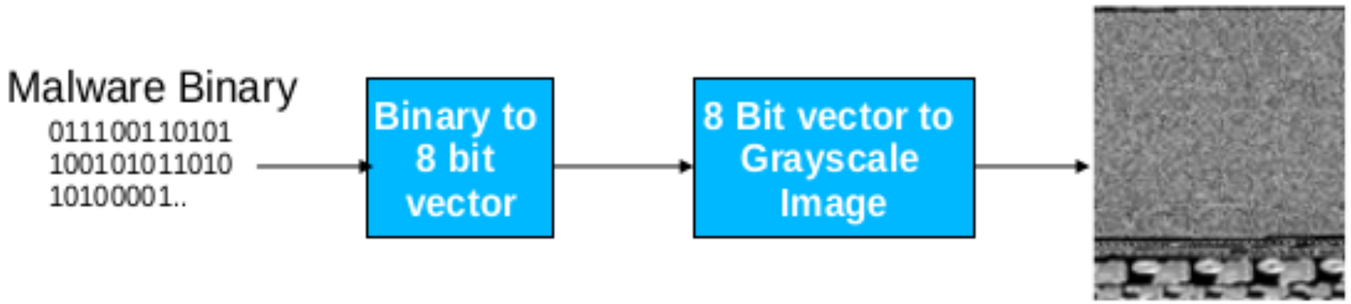
https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

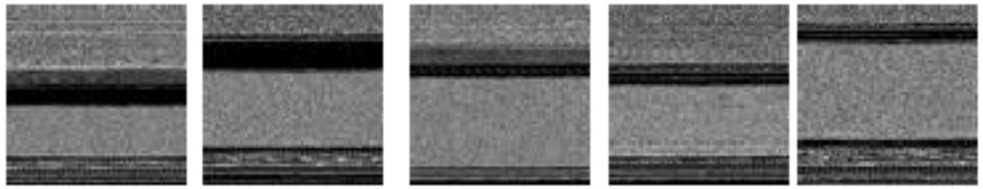
Cool CNN Applications

- Revolutionized any field having to do with images
  - Object detection/segmentation
  - Video analysis
  - Robotics
  - Self-driving cars
- Hand-writing recognition
- Heathcare
  - Breast and skin-cancer diagnosis

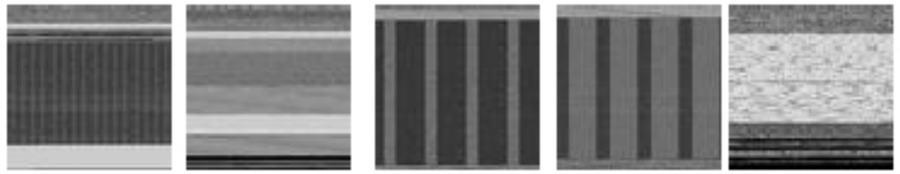
Convolution in Cybersecurity

Gibert, Daniel. "**Convolutional neural networks for malware classification**."
University Rovira i Virgili, Tarragona, Spain (2016).





Rammit Samples

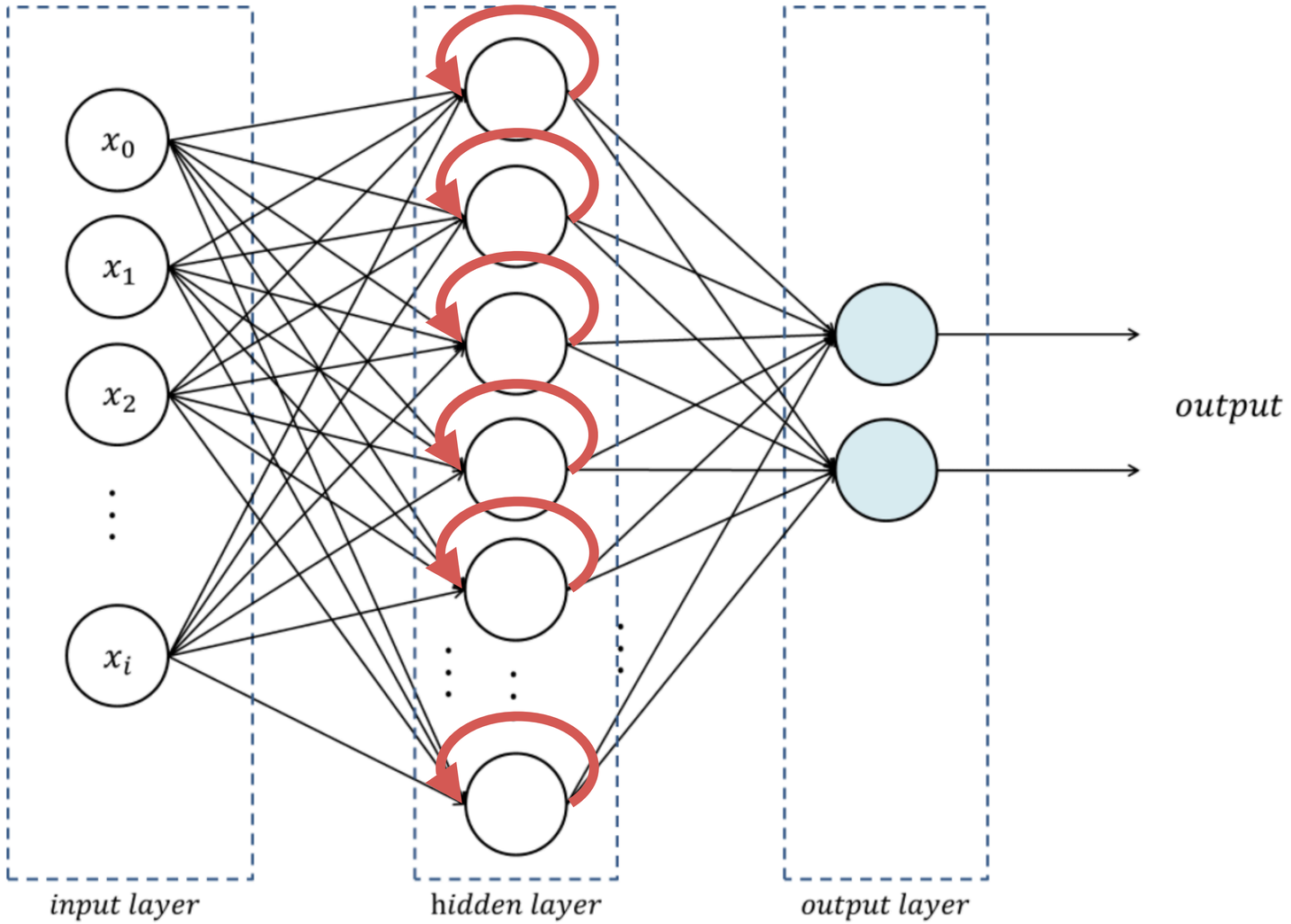


Lollipop Samples

Recurrent Neural Networks

Time-series Neural Networks

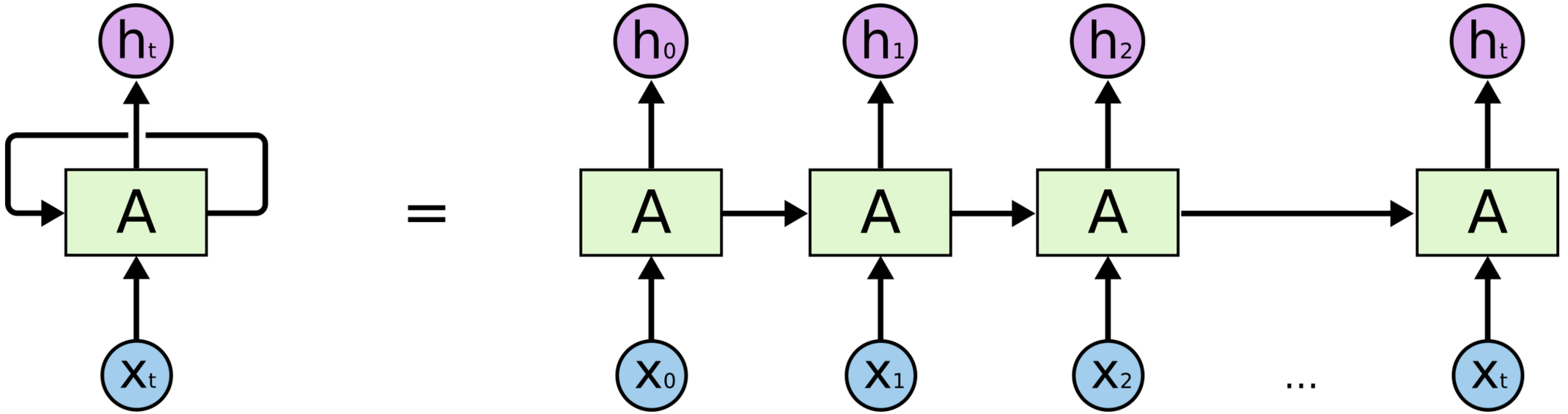
Recurrence



gtkcyber.com

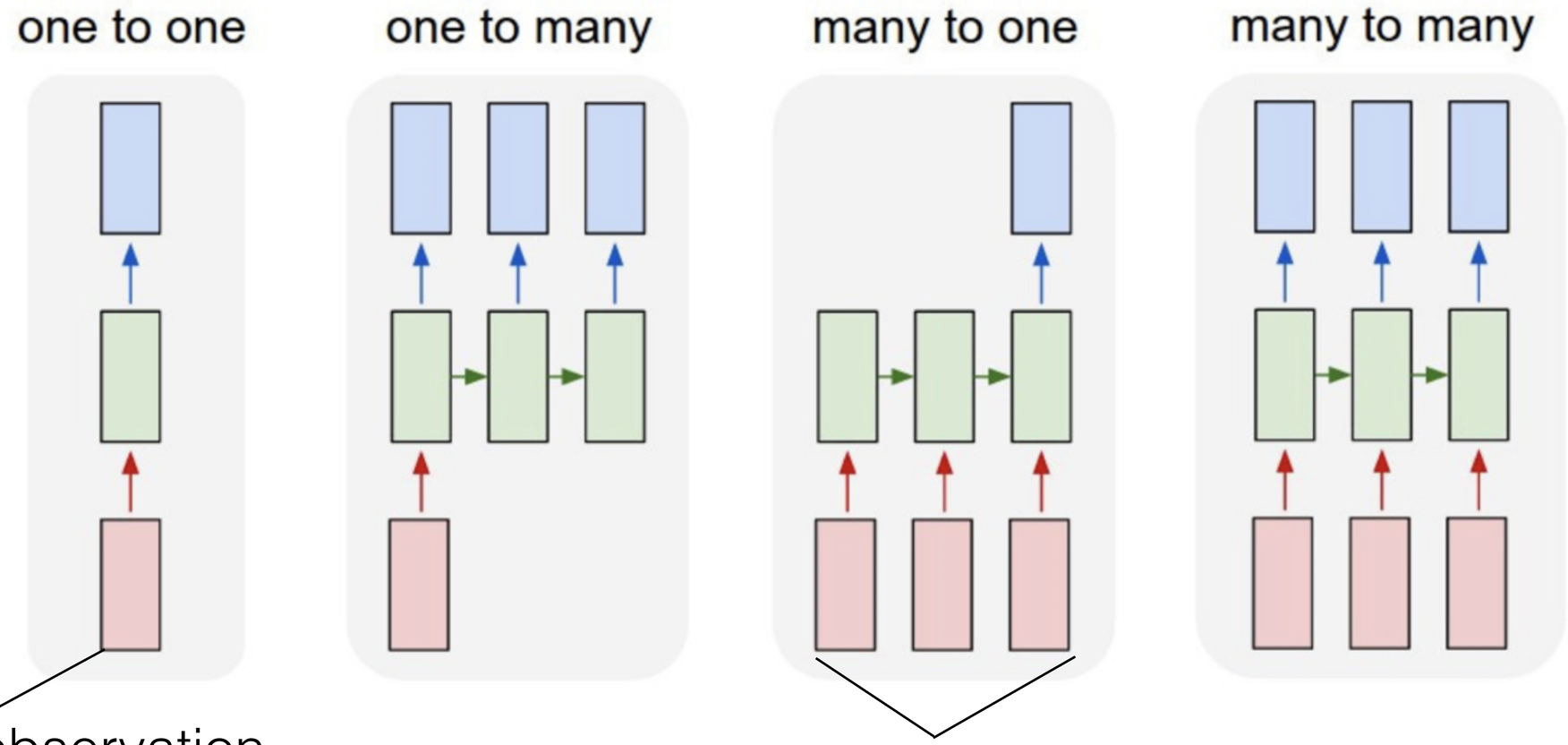
https://www.cc.gatech.edu/~san37/post/dlhc-fnn/

Recurrence - unwrapped



https://towardsdatascience.com/recurrent-neural-networks-and-lstm-4b601dd822a5

Recurrent Neural Networks (RNNs)



- 1 data observation
- Can be multidimensional

- Time series or sequence of multiple data observations
- Each can be multidimensional

RNN Variants

- Vanilla RNN ✗
- Long Short-term Memory (LSTM) ✓
- Gated Recurrent Unit (GRU) ✓

Cool RNN Applications

- Natural Language Processing (NLP)
  - Machine Translation (Google's seq2seq)
- Image Captioning
- Webpage Classification
- Domain Generation Algorithm (DGA) Classification

RNNs in Cybersecurity

Torres, Pablo, et al. "**An analysis of Recurrent Neural Networks for Botnet detection behavior**." *2016 IEEE biennial congress of Argentina (ARGENCON)*. IEEE, 2016.

TABLE I
SYMBOL ASSIGNMENT STRATEGY FOR BUILDING BEHAVIORAL MODEL ACCORDING TO THE STRATOSPHERE PROJECT.

	Size Small			**Size Medium**			**Size Large**		
	Dur. Short	**Dur. Med**	**Dur. Long**	**Dur. Short**	**Dur. Med**	**Dur. Long**	**Dur. Short**	**Dur. Med**	**Dur. Long**
Strong Per	a	b	c	d	e	f	g	h	i
Weak Per.	A	B	C	D	E	F	G	H	I
Weak Non-Per.	r	s	t	u	v	w	x	y	z
Strong Non-Per	R	S	T	U	V	W	X	Y	Z
No Data	1	2	3	4	5	6	7	8	9

Symbols for time difference	
Between 0 and 5 seconds:	.
Between 5 and 60 seconds:	,
Between 60 and 5 mins:	+
Between 5 mins and 1 hour	*
Timeout of 1 hour:	0

2.4.R*R.R.R*a*b*a*a*b*b*a*R.R*R.R*a*a*b*a*a*a*a*

Fig. 1. An example of the behavioral model of connection from IP address 10.0.2.103 to destination port 53 at IP address 8.8.8.8 port 53 using protocol UDP.

Deep Learning Tools

- Tensorflow (Google)
  - The most math-heavy but most versatile
- PyTorch (Facebook)
- Caffe (UC Berkeley)
- Keras
  - Abstraction to Tensorflow to make it easier
  - Probably use this

Deep Learning Takeaways

- Deep learning isn't magic
  - The size of our datasets is what's really magic
- Certain problems and problem areas are perfect for deep learning
  - Others are not
- Deep learning is hard to implement
  - But the tools are getting easier every day
- RNNs probably have the most application in Cybersecurity