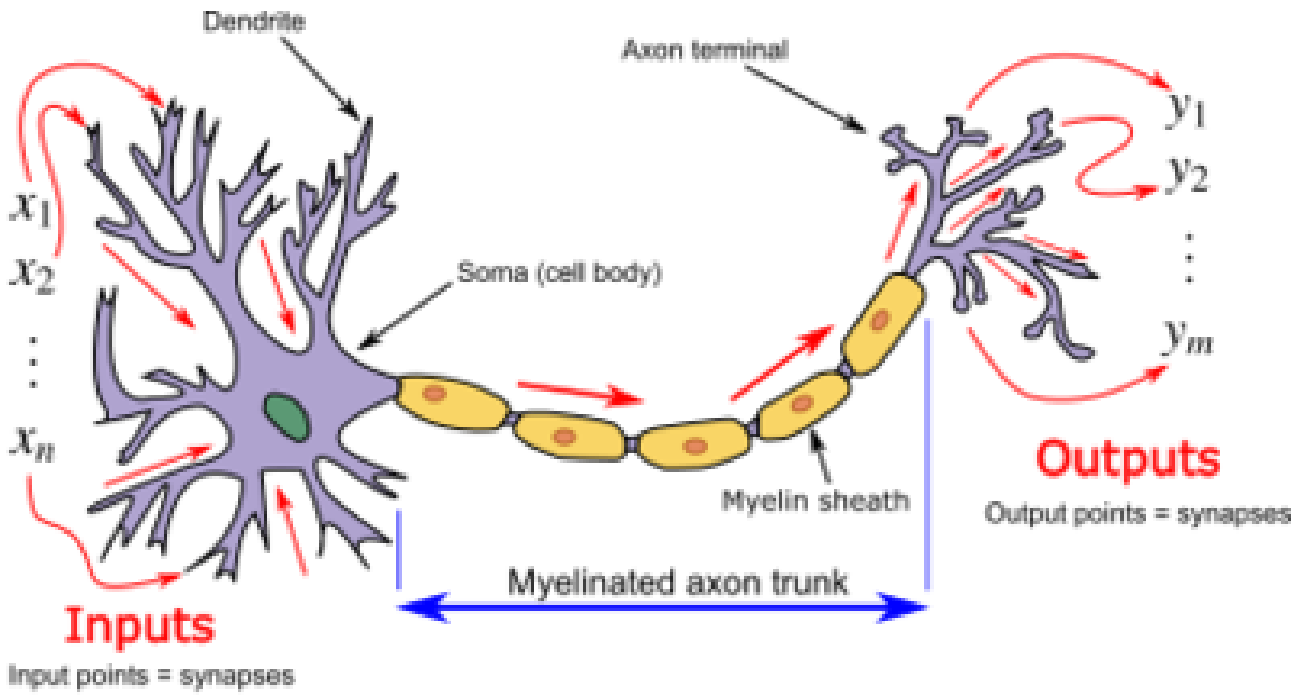
Neural Networks

CNN4N Journal Club

Amr Elsawy

Biological Neuron

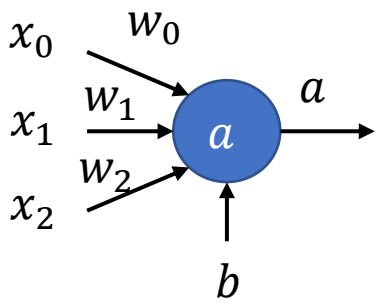


https://en.wikipedia.org/wiki/Biological_neuron_model

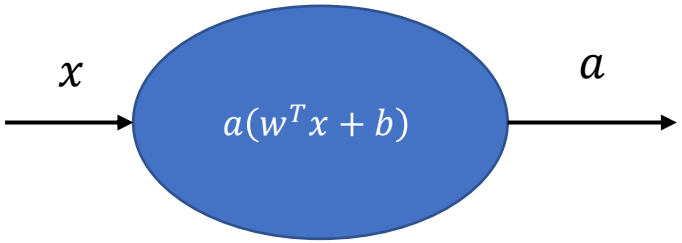
Artificial Neural Networks (ANN)

Artificial Neuron

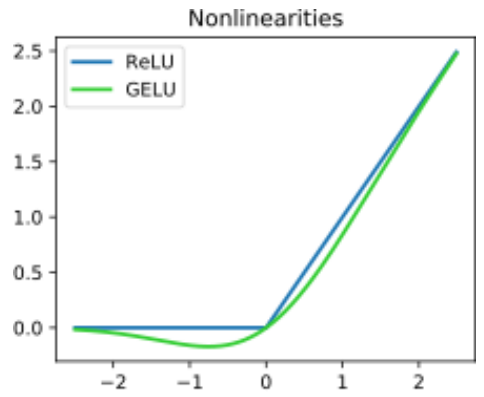
- It is a biologically inspired computational unit.
- It is the basic (i.e., primary) building unit in artificial neural networks (ANN).
- It applies two operations: linear (i.e., dot product) and non-linear (i.e., activation).



$$z = w^T x + b$$
$$a(z)$$



How many unknows?
What are they called?

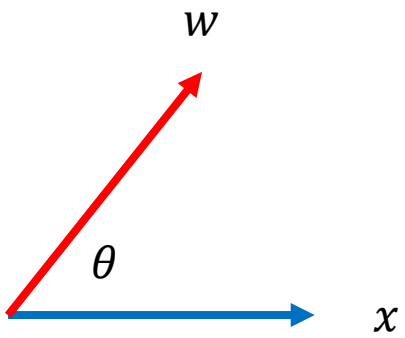


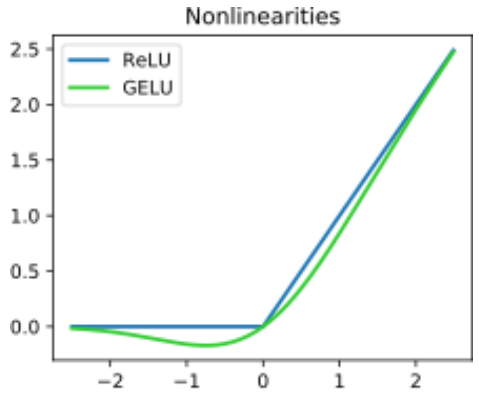
https://en.wikipedia.org/wiki/Rectifier_(neural_networks)

Cont.

- Assume unit vectors, what does $w^T x$ represent?

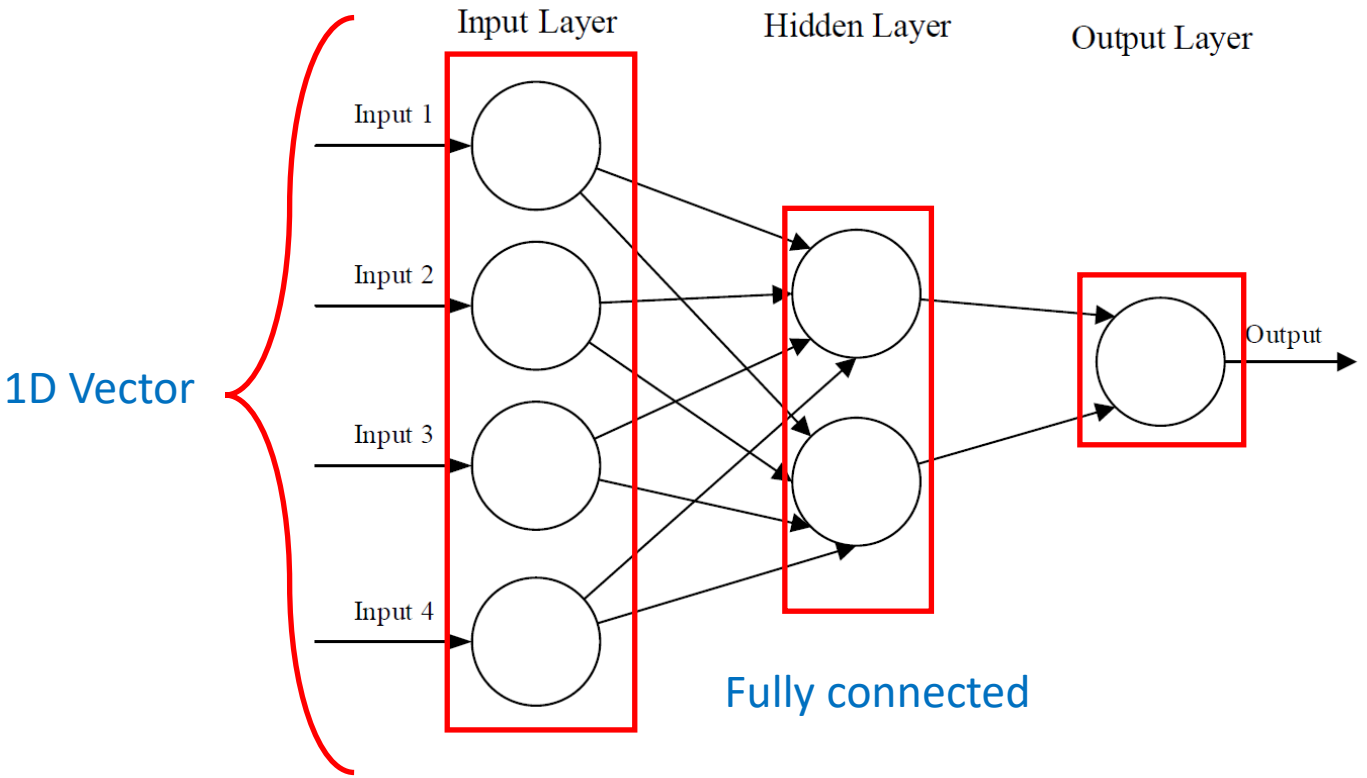
- What does the non-linearity do?





https://en.wikipedia.org/wiki/Rectifier_(neural_networks)

Feed Forward Network (FNN)



How many neurons in the figure?

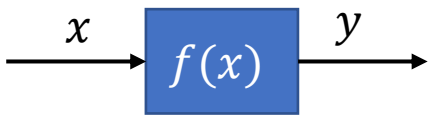
?

Computations flow into one direction

There is no spatial information of the input

Artificial Neural Networks (ANN)

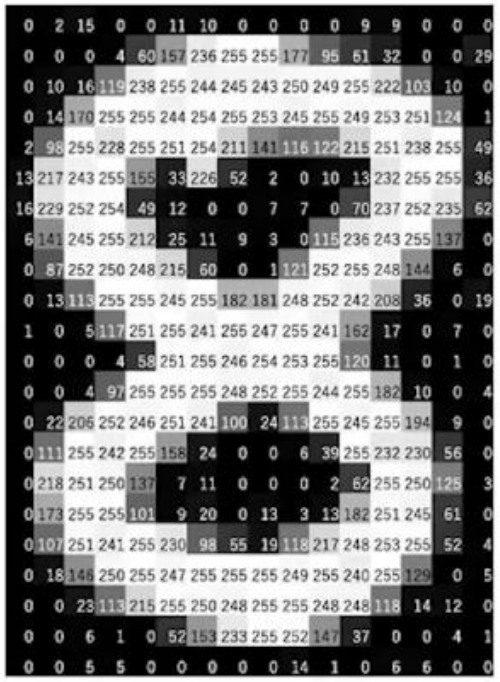
- ANN is a mathematical model.
- ANN has some parameters that defines its complexity (i.e., dimensionality).
- Deep ANN has more than one hidden layer.
- Learning is done by finding the best values for these parameters.

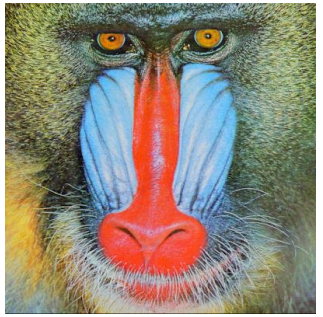


Network

Image Data

- Image is represented as a 2D (i.e., grayscale) or 3D matrix (i.e., color).





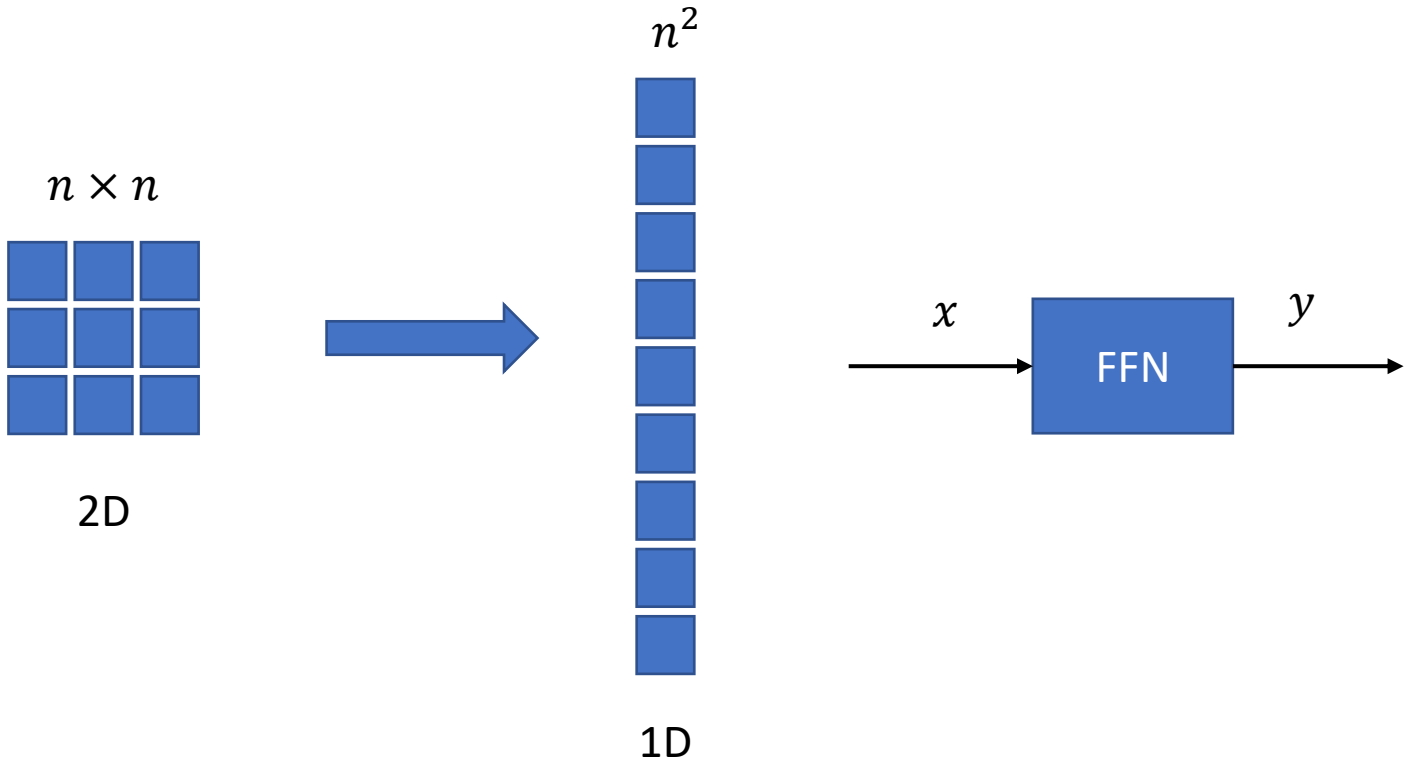
?

https://www.r-bloggers.com/2016/01/color-quantization-in-r/

https://www.analyticsvidhya.com/blog/2021/03/grayscale-and-rgb-format-for-storing-images/

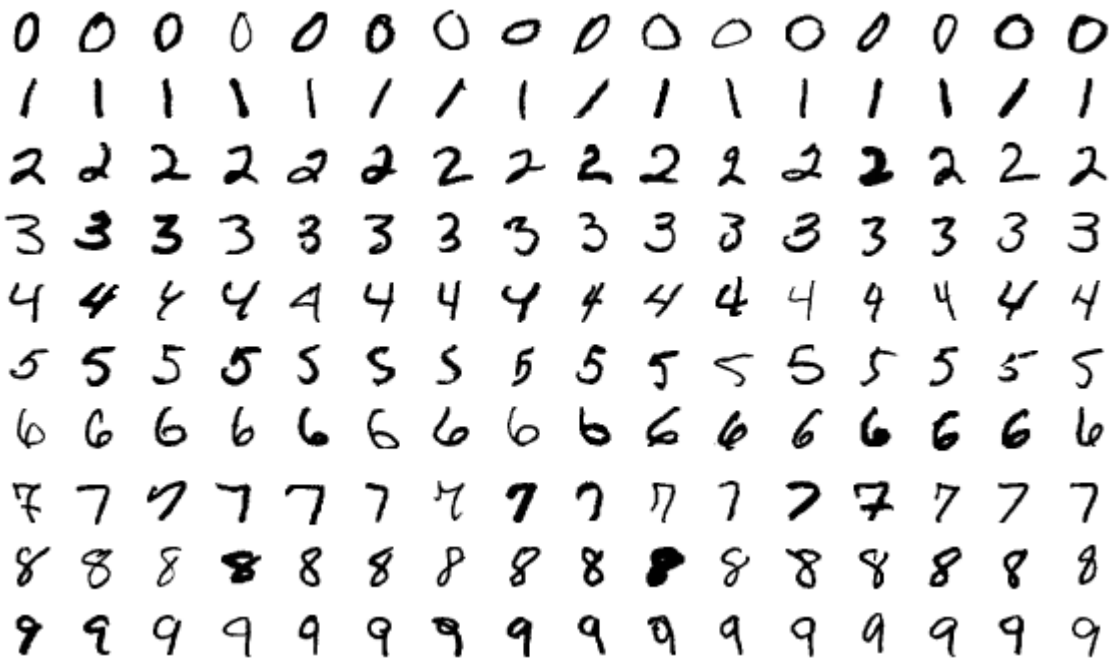
Cont.

- To deal with 2D image data, it is converted into 1D vector.



Cont.

- Benchmarking datasets such as the MNIST database of handwritten digits are suitable for most forms of ANN.



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

Cont.

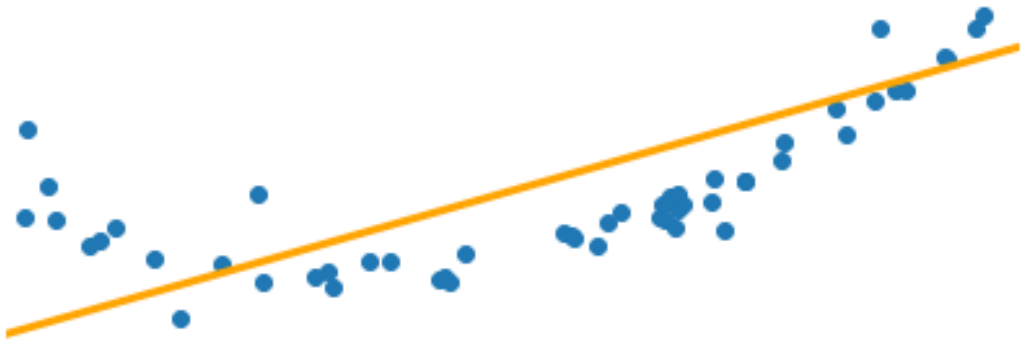
- Images in MNIST dataset have dimensionality of $28 \times 28 = 784$ pixels.
- Also, the values are black and white (i.e., no colors).
- What if images are colored and have dimensionality of 64×64?

Best Model

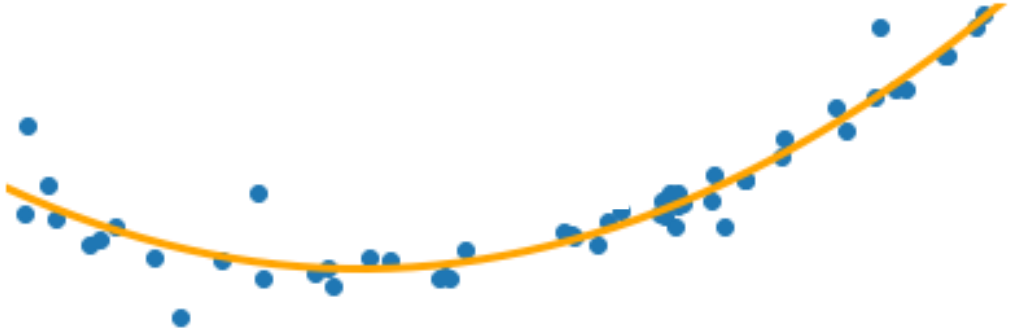
What is the best polynomial that fits the points? How do we measure that? Why?

?

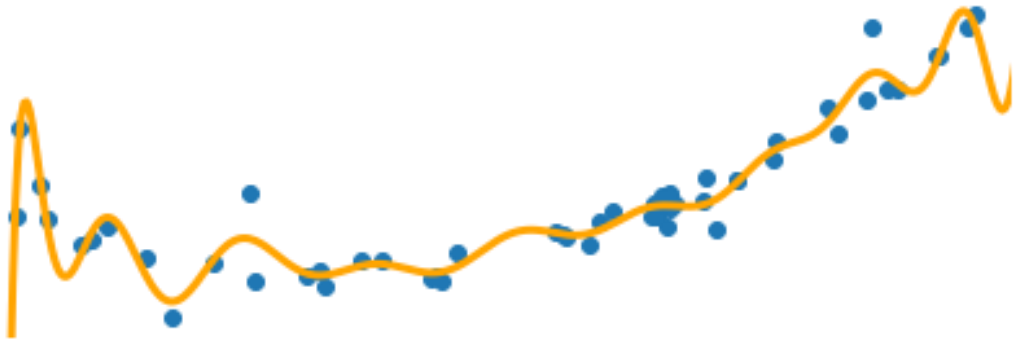
1st order



2nd order



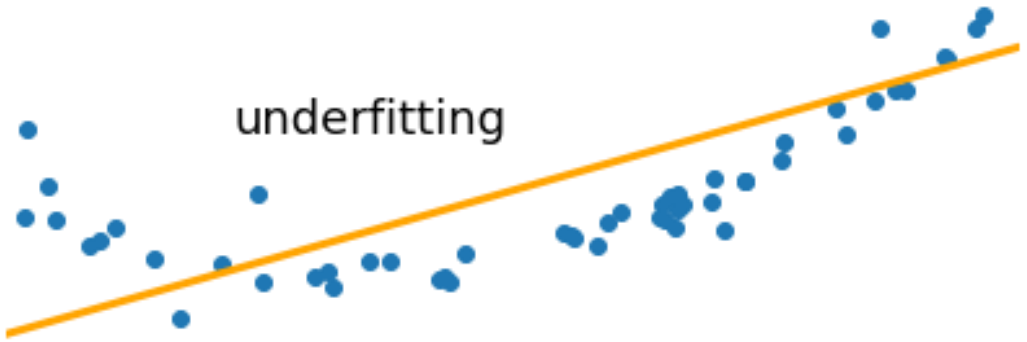
Very high order



https://bookdown.org/gmli64/do_a_data_science_project_in_10_days/models-underfitting-and-overfitting.html

Overfitting

1st order



2nd order

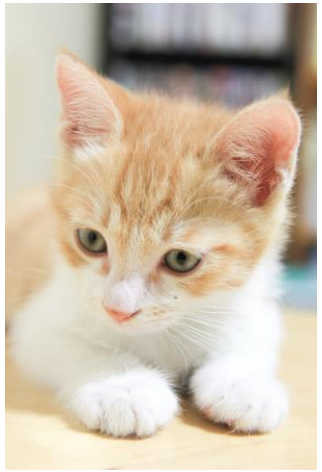
Very high order

https://bookdown.org/gmli64/do_a_data_science_project_in_10_days/models-underfitting-and-overfitting.html

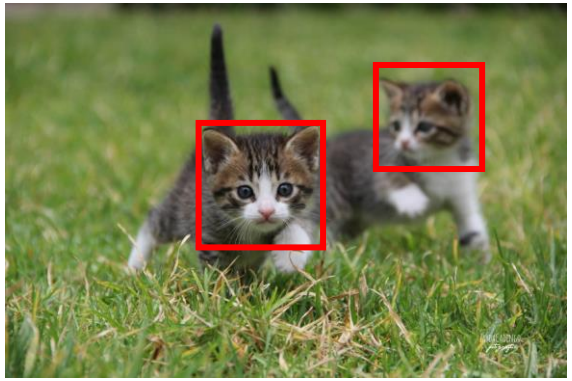
"One of the largest limitations of traditional forms of ANN is that they tend to struggle with the computational complexity required to compute image data."

Data

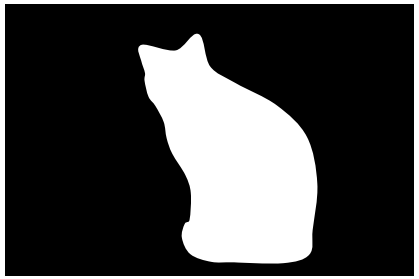
- Input data can be numeric, text, image, etc.
- Output data can be binary labels, categorical labels, detection, segmentation, annotations, etc.



Cat



Detection



Segmentation

A cat is sleeping

Learning

- Learning is done by training the network using the training data.
- The network should learn how to predict the output labels from the input data.

- There are many types of learning: supervised, unsupervised, semi-supervised, and self-supervised.

supervised	Classification
Unsupervised	Clustering
Semi-supervised	Automatic labelling
Self-supervised	Autoencoders

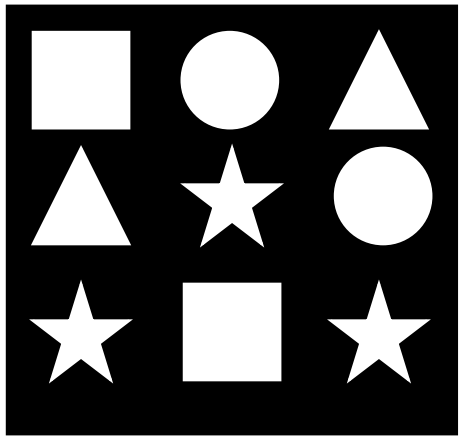
- In **supervised** learning, the training labels are known.

- In **unsupervised** learning, the training labels are unknown.

Convolutional Neural Networks (CNN)

Convolution (*)

Image



*

Kernel 1

Feature map 1

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

*

Kernel 2

Feature map 2

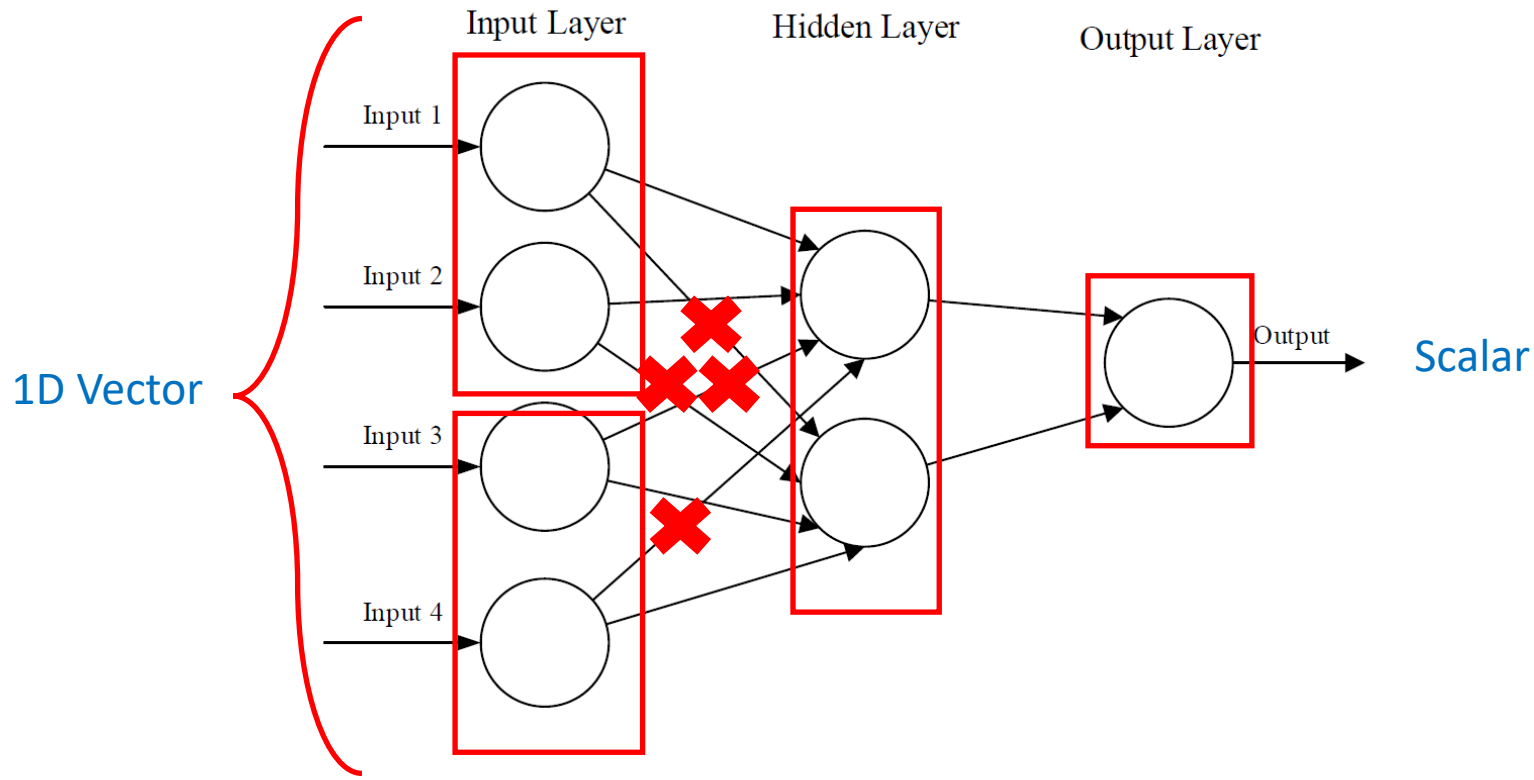
$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Assume 0/1 matching output for simplicity

Local not global matching

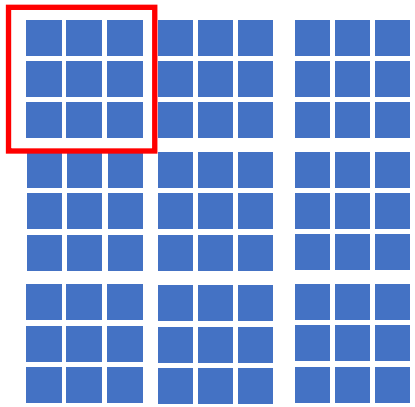
Revisit ANN

Imagine we split input …



Cont.

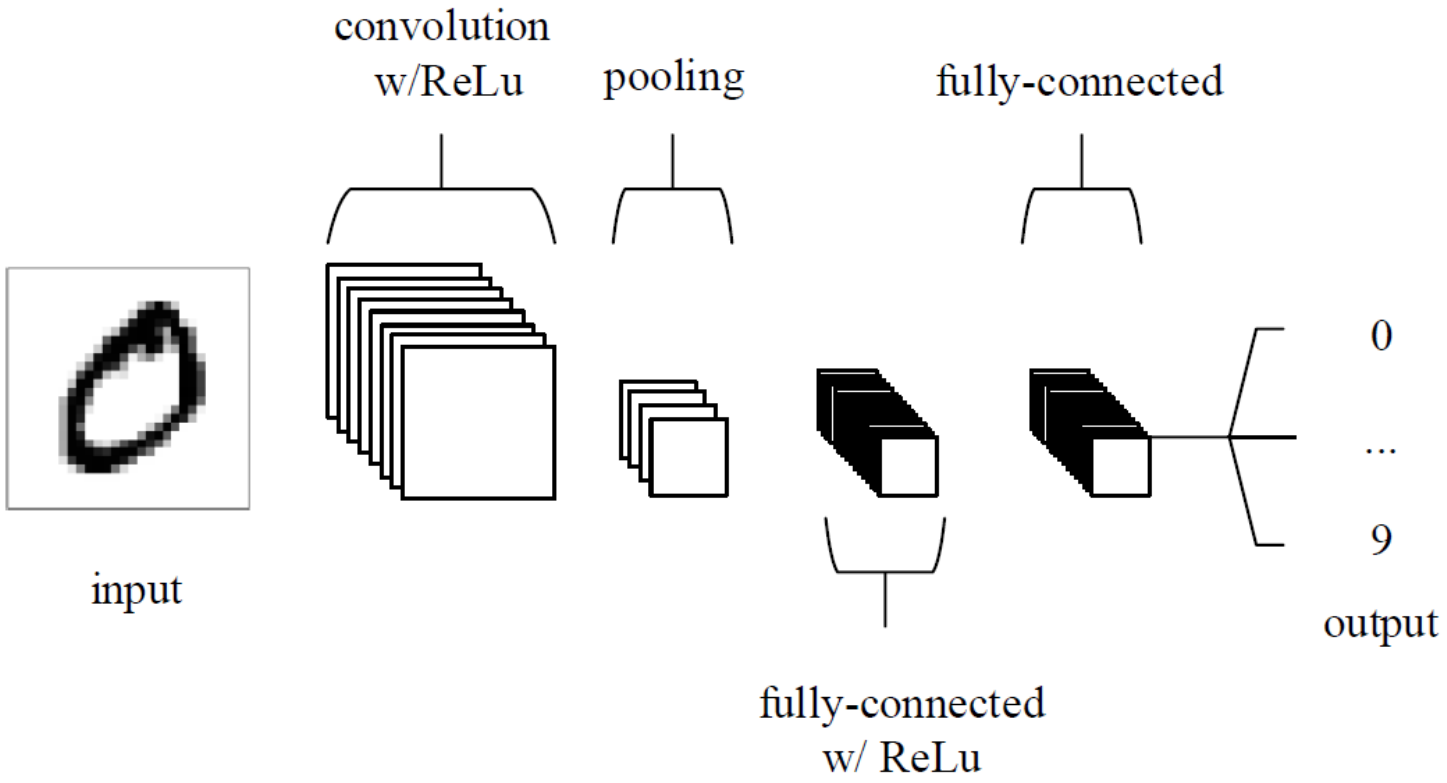
- Unlike standard ANNS, the neurons within any given layer will only connect to a small region of the layer preceding it.
- Another difference is that they are shared.



Cont.

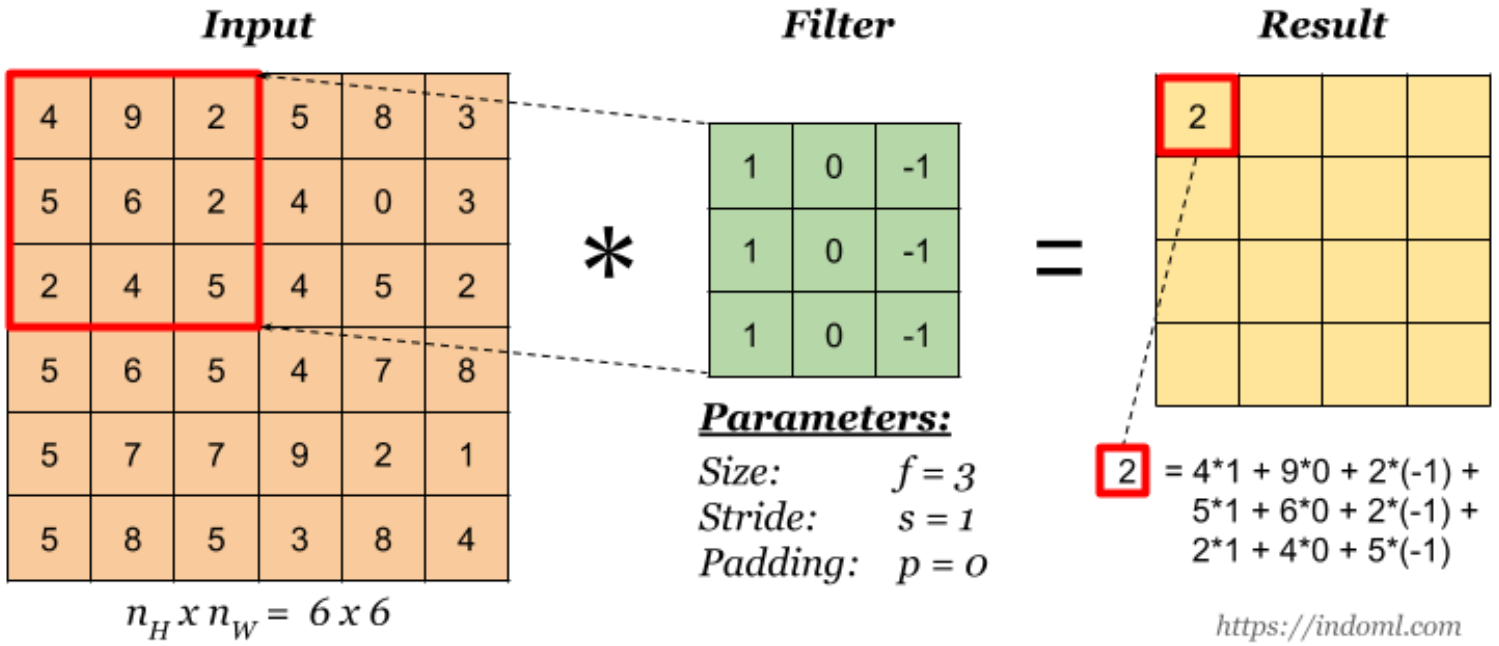
- CNNs are primarily used in the field of computer vision to encode image-specific features into the architecture, which makes the network more suited for image-focused tasks - whilst further reducing the parameters required to set up the model.

CNN Architecture

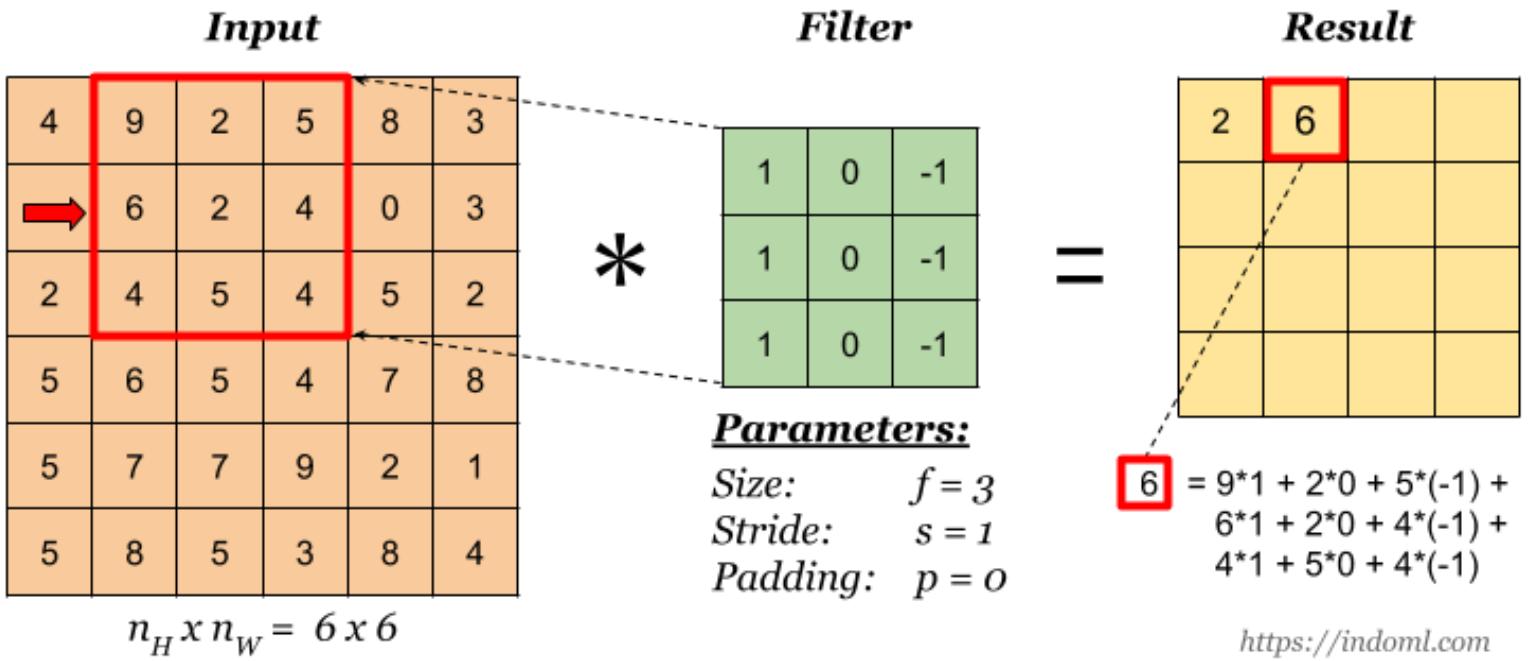


Simple CNN architecture, comprised of just five layers

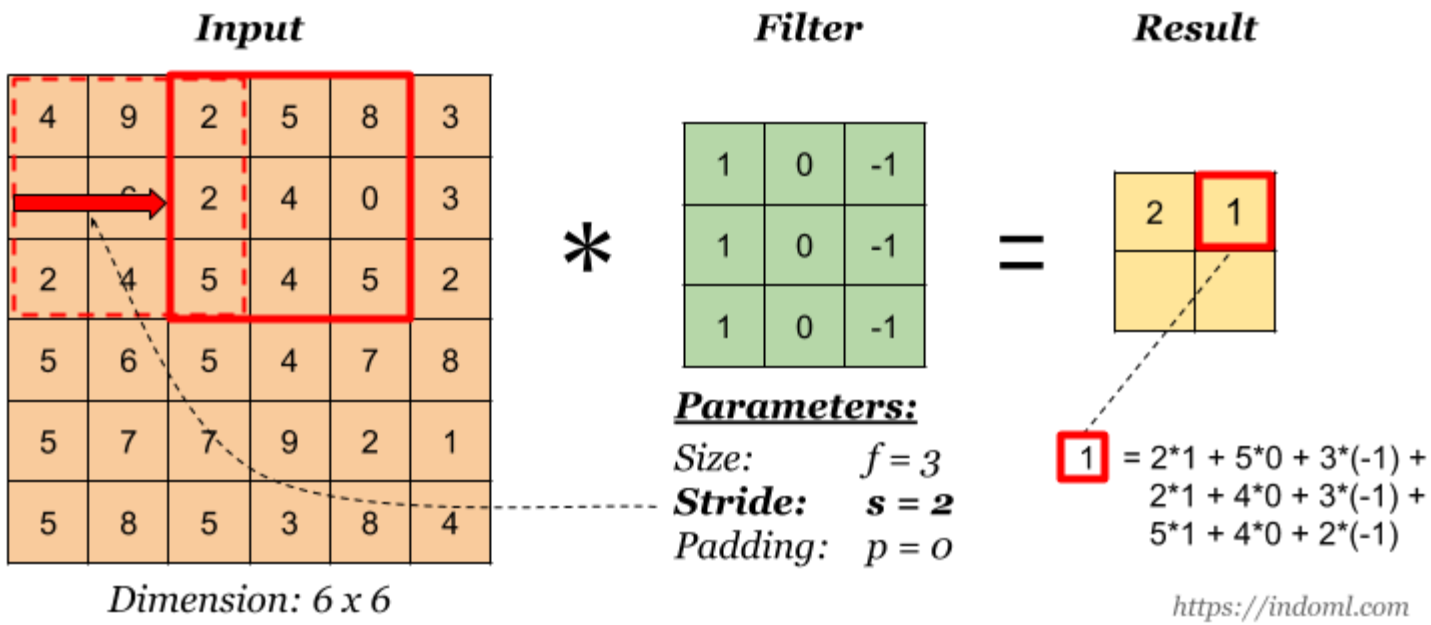
Basic Convolution Operation



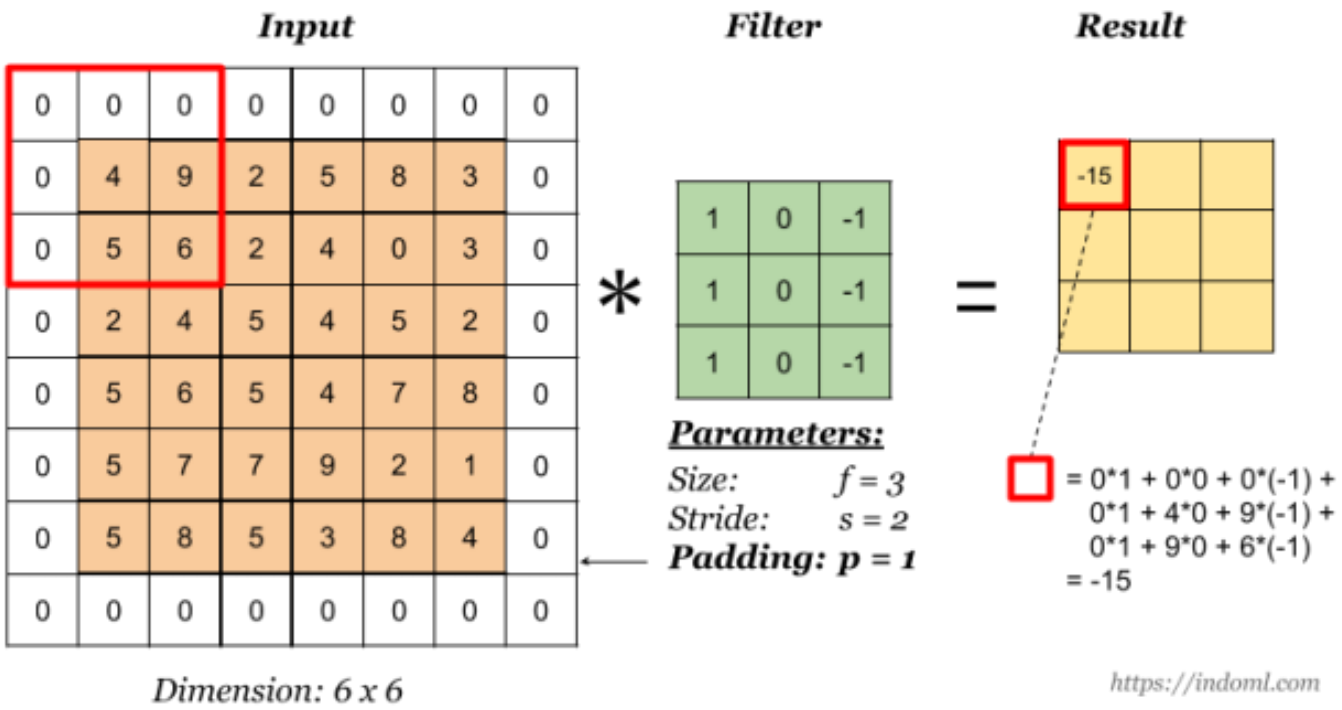
Cont.


Input
Filter
Result
4 9 2 5 8 3
6 2 4 0 3
2 4 5 4 5 2
5 6 5 4 7 8
5 7 7 9 2 1
5 8 5 3 8 4
1 0 -1
1 0 -1
1 0 -1
2 6
Parameters:
Size: f = 3
Stride: s = 1
Padding: p = 0
6 = 9*1 + 2*0 + 5*(-1) +
6*1 + 2*0 + 4*(-1) +
4*1 + 5*0 + 4*(-1)
n_H x n_W = 6 x 6
https://indoml.com

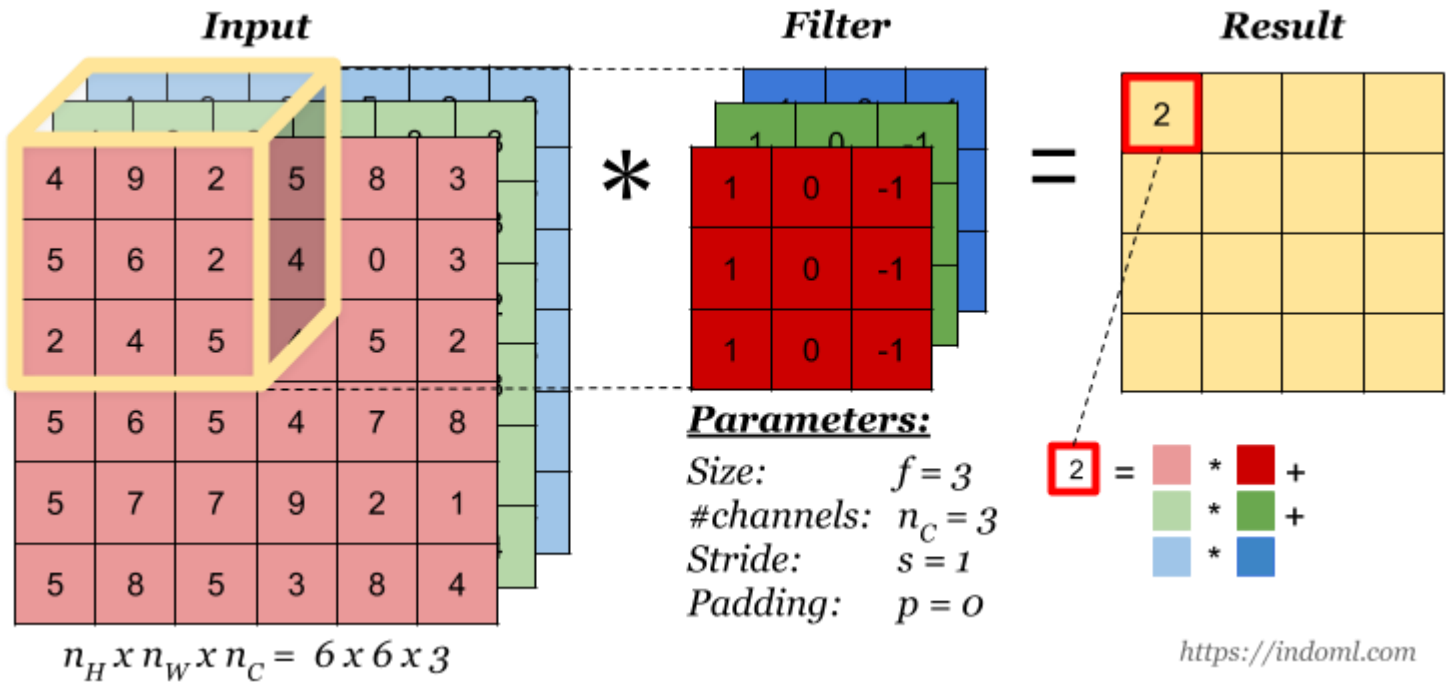
Stride


Input
Filter
Result
4 9 2 5 8 3
2 4 0 3
2 4 5 4 5 2
5 6 5 4 7 8
5 7 7 9 2 1
5 8 5 3 8 4
1 0 -1
1 0 -1
1 0 -1
2 1
Parameters:
Size: f = 3
Stride: s = 2
Padding: p = 0
1 = 2*1 + 5*0 + 3*(-1) +
2*1 + 4*0 + 3*(-1) +
5*1 + 4*0 + 2*(-1)
Dimension: 6 x 6
https://indoml.com

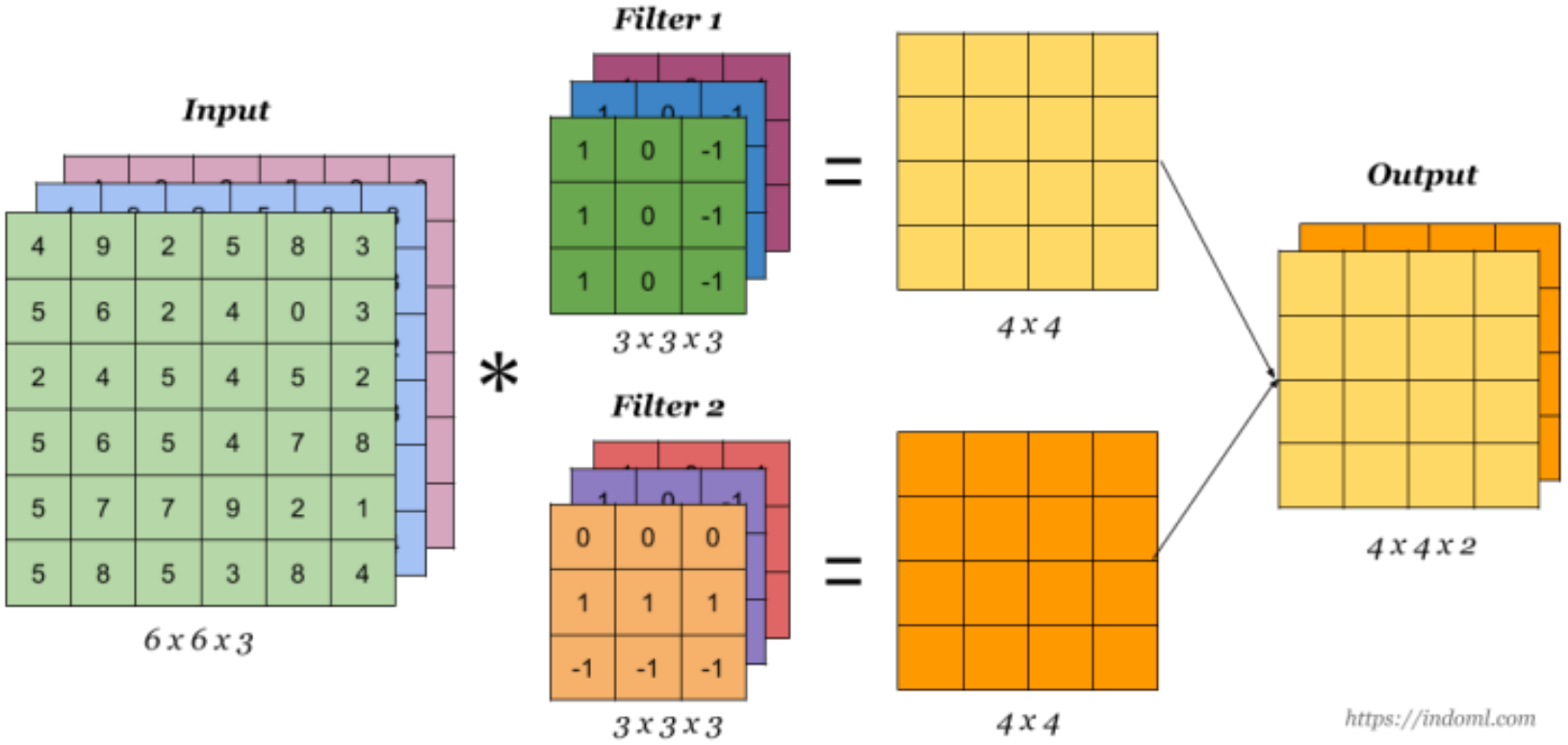
Padding


Input
Filter
Result
0 0 0 0 0 0 0 0
0 4 9 2 5 8 3 0
0 5 6 2 4 0 3 0
0 2 4 5 4 5 2 0
0 5 6 5 4 7 8 0
0 5 7 7 9 2 1 0
0 5 8 5 3 8 4 0
0 0 0 0 0 0 0 0
1 0 -1
1 0 -1
1 0 -1
-15
Parameters:
Size: f = 3
Stride: s = 2
Padding: p = 1
= 0*1 + 0*0 + 0*(-1) +
0*1 + 4*0 + 9*(-1) +
0*1 + 9*0 + 6*(-1)
= -15
Dimension: 6 x 6
https://indoml.com

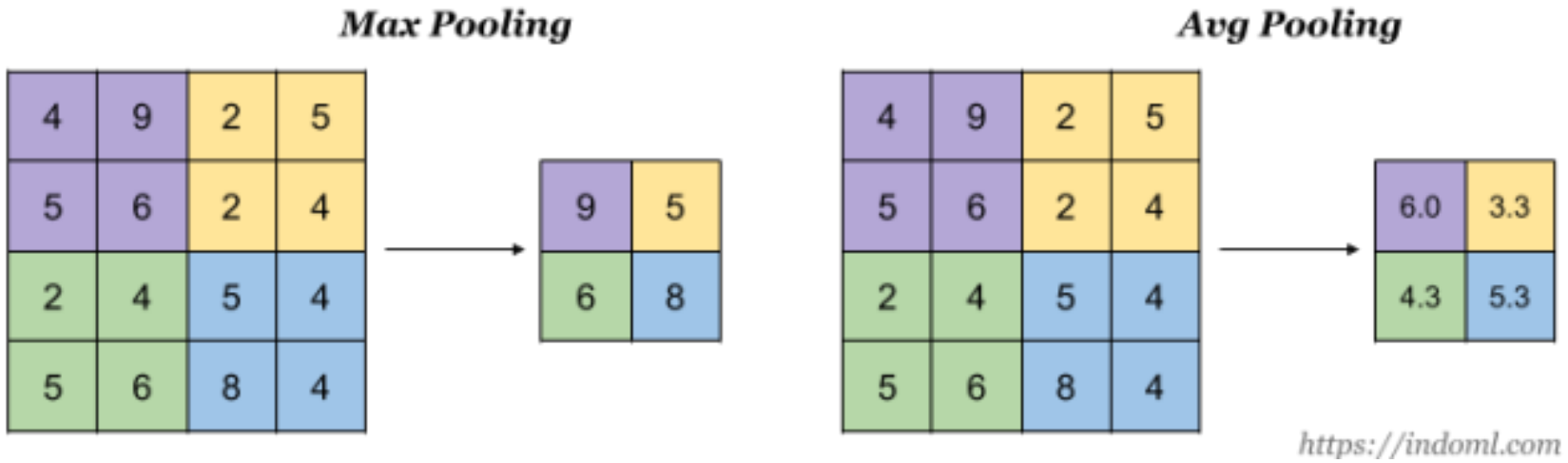
Convolution Operation on Volume



Convolution Operation with Multiple Filters



Pooling Layer

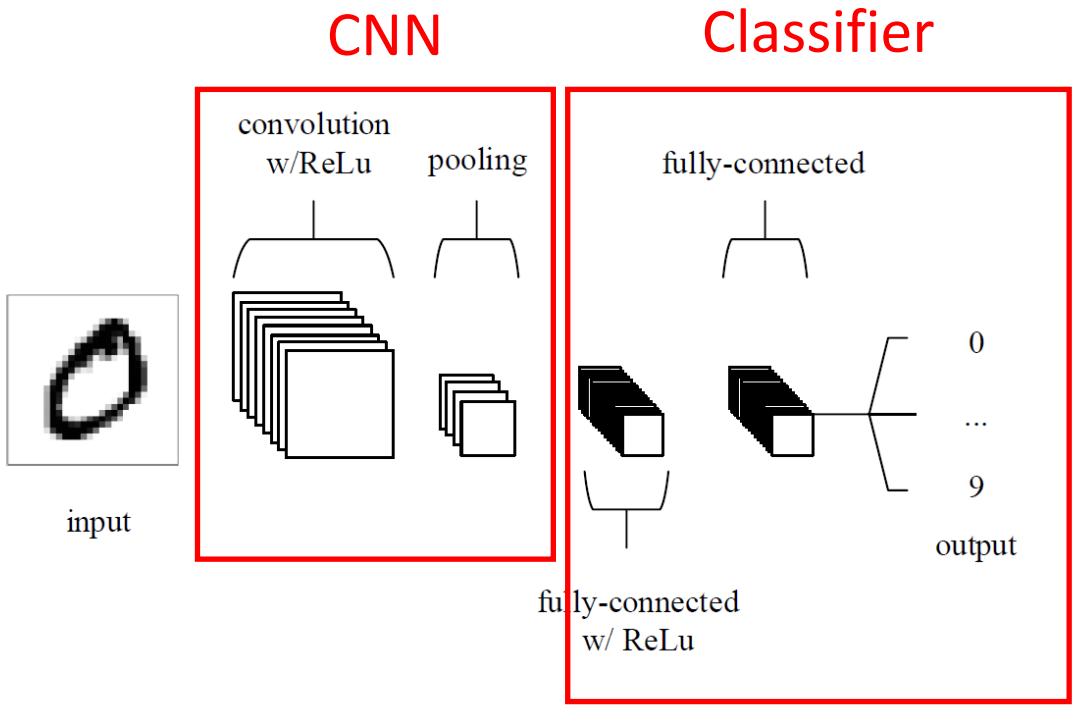


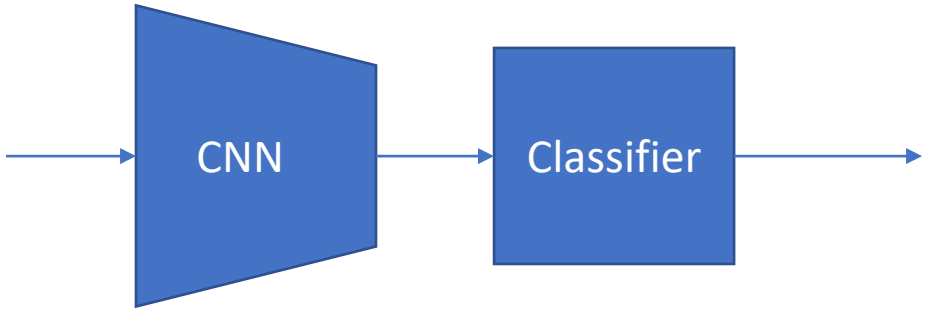
Fully Connected Layer

- It is like any hidden layer in ANN.
- It has complete (i.e., global) connections to its inputs.

Deep Network …

- Simple deep network consists of
  - Convolutional backbone
  - Classifier

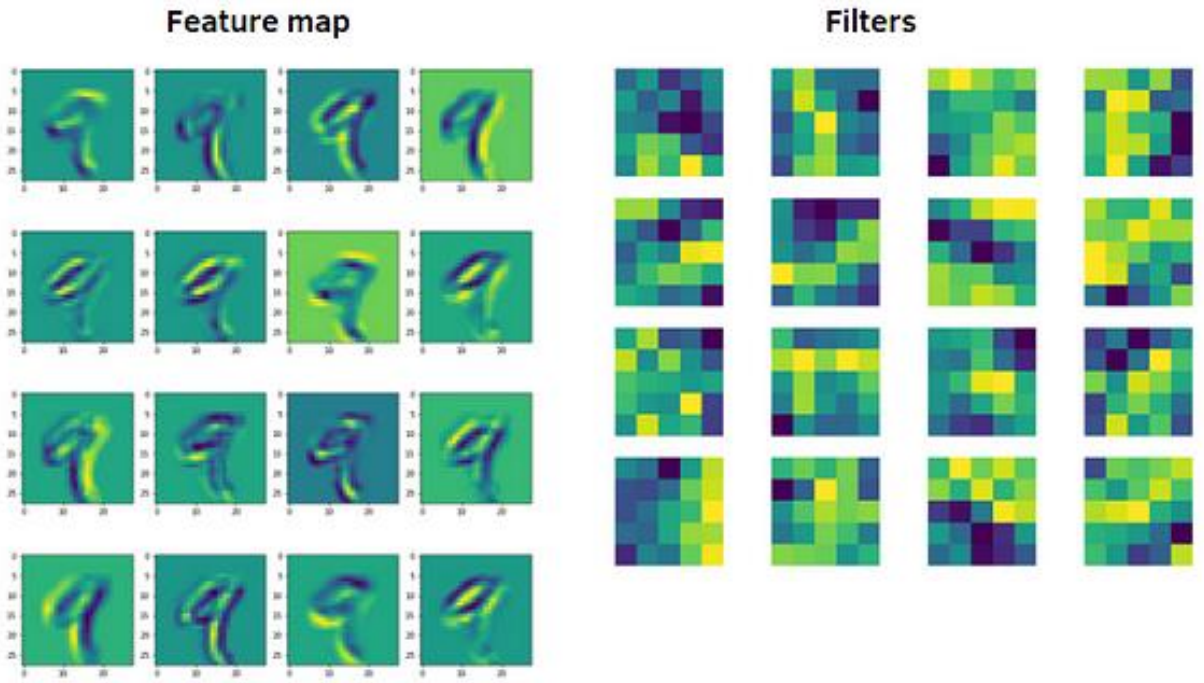




What does CNN learn?

- It learns …

?



https://medium.com/dataseries/visualizing-the-feature-maps-and-filters-by-convolutional-neural-networks-e1462340518e

Thanks ☺