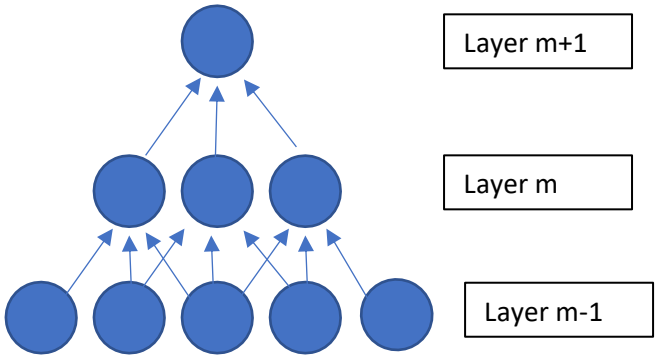
Convolutional Neural Networks

Convolutional neural networks (CNNs) are typically used to recognize patterns, features present in images but they are also used for spatial data analysis, computer vision, NLP, signal processing, etc.

The most important function or operation in CNN is convolution.

Convolution:

Convolution is an orderly procedure where two sources of information are intertwined, it is an operation that changes a function into something else.



What CNNs actually do?

CNNs make use of filters also known as kernels, to detect features such as edges, present throughout an image.

CNNs have 4 main operations:

- Convolution
- Non-linearity (ReLU)
- Pooling or sub-sampling
- Classification (fully connected layers)

The first layer is always convolutional layer, that applies convolution operation to input, passing the result to the next layer. A convolution converts all the pixels in its receptive field into a single value. The final output of the convolutional layer is a vector.

If we add kernel of size *ks* by *ks*(with ks an odd number), the necessary padding on each side to keep the same shape is ks//2. An even number for ks would require a different amount of padding on top/bottom and left/right, but in practice, we never use an even filter size.

We can move over two pixels after each kernel application. This is called stride-2 convolution. In an image of size h by w, using a padding of 1 and stride-2 will give us result of size (h+1)//2 by (w+1)//2.

Simple CNN:

Simple cnn can be:

```
simple_cnn = sequential(
conv(1,4),          # 14 x 14
conv(4,8),          # 7 x 7
conv(8,16),         # 4 x 4
conv(16,32)    ,    # 2 x 2
conv(32,2,act=False),      #1 x 1
Flatten()
)
```

The first digit in conv(x,y) is input channel, and the second digit is output channel.

Input channel:

Input to convolutional layer typically has multiple channels, especially for color images. For example, a color image with RGB has 3 input channels.

Output channels:

They are the number of feature images or channels produced as result of applying convolution operation to the input. Each output channel in convolutional layer represents a different feature or filter learned by the neural network. These features could capture edges, textures, etc.

You can think of each output channel as an image in itself but with different features highlighted. These channels are generated through the convolution operation, where a small filter or kernel slides over input, computing a weighted sum of values at each location to produce the output.