upGrad

PLACEMENT REFRESHER PROGRAM

Session 13 - Deep Learning 1

ANN & CNN

By
Ritesh Kumar Pandey

Agenda

- Deep Learning
- ANN
- CNN

MCQ / MSQ

upGrad

Machine learning is a subset of which of the following.

A) Deep Learning
B) Reinforcement Learning
C) Artificial Intelligence
D) All of the above

MCQ / MSQ

upGrad

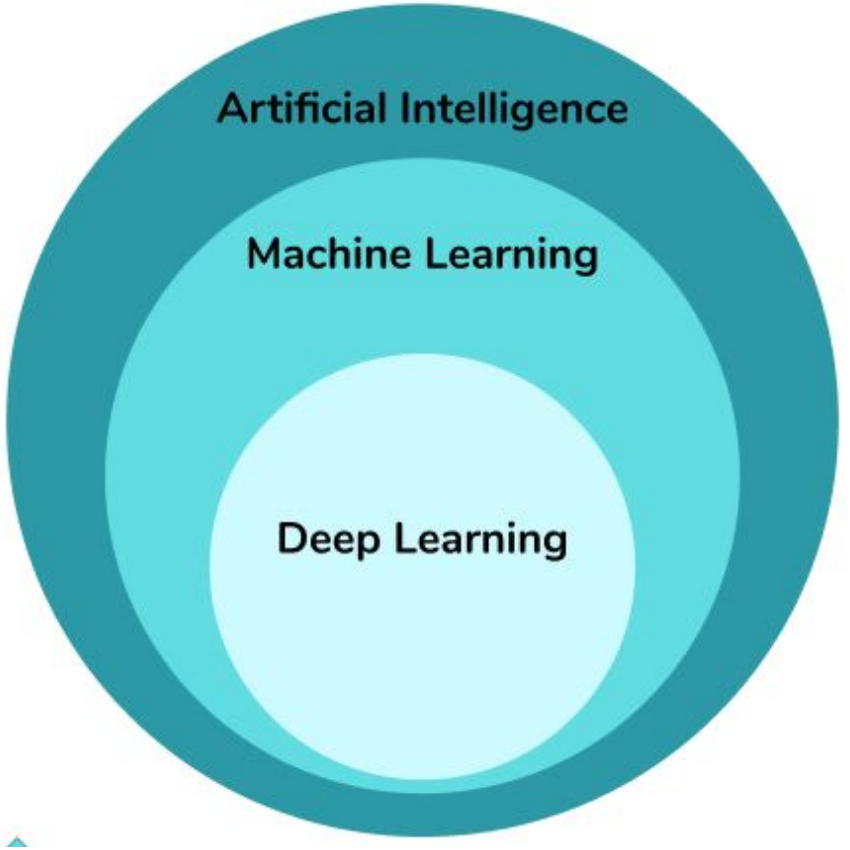
Machine learning is a subset of which of the following.

A) Deep Learning
B) Reinforcement Learning
C) Artificial Intelligence
D) All of the above

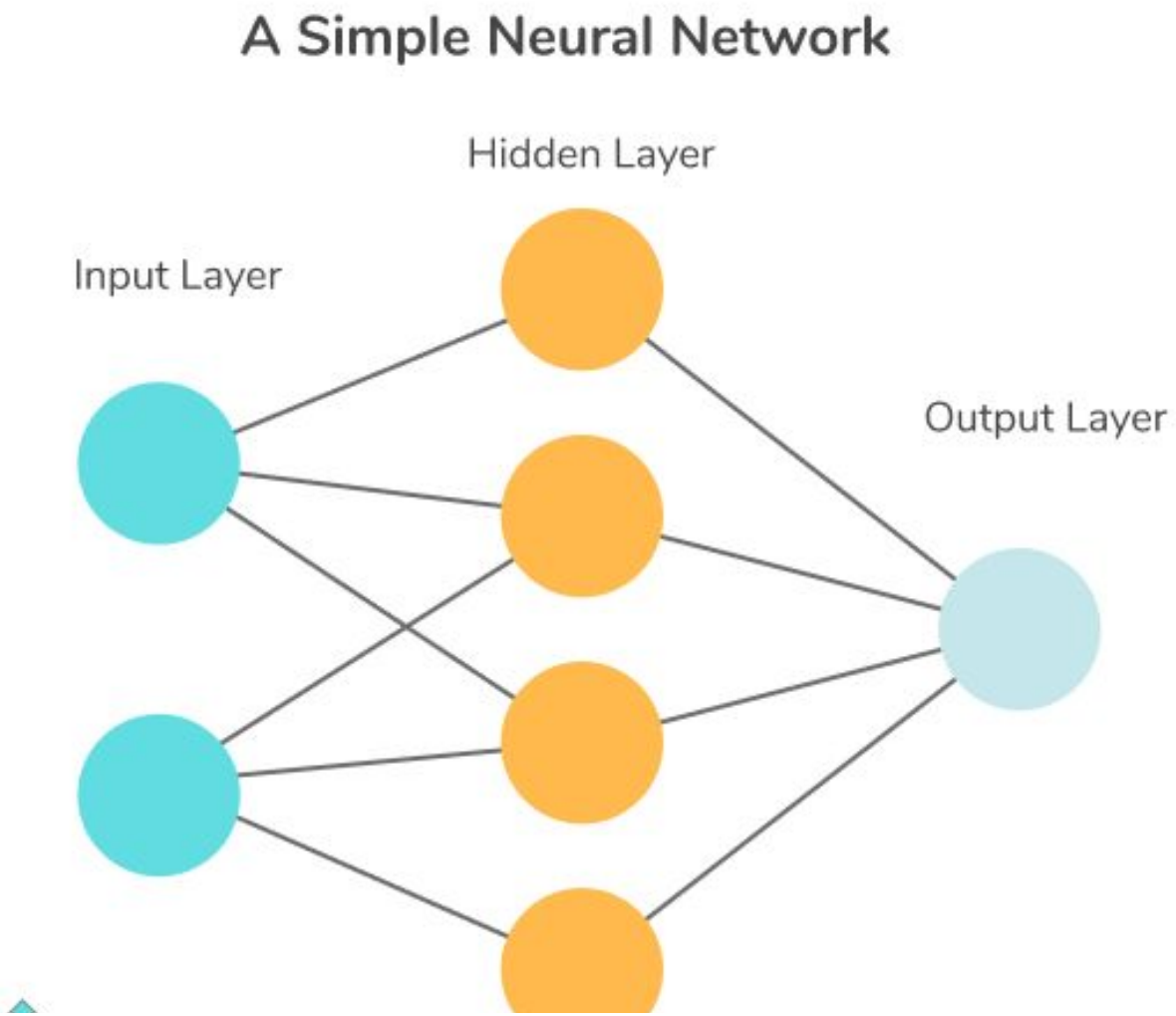
Deep Learning

upGrad

Deep learning is a subset of machine learning that is entirely based on artificial neural networks. Because neural networks are designed to mimic the human brain, deep learning is likewise a human brain mimic. We don't have to explicitly program everything in deep learning. We train a model on a training dataset and improvise it until the model predicts almost correctly on the testing and validation dataset as well.



Neural Networks are artificial systems that have a lot of resemblance to the biological neural networks in the human body. A neural network is a set of algorithms that attempts to recognize underlying relationships in a batch of data using a method that mimics how the human brain works.



MCQ / MSQ

upGrad

What is an activation value?

a) weighted sum of inputs
b) threshold value
c) main input to neuron
d) none of the mentioned

MCQ / MSQ

upGrad

What is an activation value?

a) weighted sum of inputs
b) threshold value
c) main input to neuron
d) none of the mentioned

MCQ / MSQ

upGrad

The process of adjusting the weight is known as?

a) activation
b) synchronisation
c) learning
d) none of the mentioned

MCQ / MSQ

upGrad

The process of adjusting the weight is known as?

a) activation
b) synchronisation
c) learning
d) none of the mentioned

MCQ / MSQ

upGrad

What is generalization?

a) the ability of a pattern recognition system to approximate the desired output values for pattern vectors which are not in the test set.
b) the ability of a pattern recognition system to approximate the desired output values for pattern vectors which are not in the training set.
c) can be either way
d) none of the mentioned

MCQ / MSQ

upGrad

What is generalization?

a) the ability of a pattern recognition system to approximate the desired output values for pattern vectors which are not in the test set.
b) the ability of a pattern recognition system to approximate the desired output values for pattern vectors which are not in the training set.
c) can be either way
d) none of the mentioned

MCQ / MSQ

upGrad

How many layers Deep learning algorithms are constructed?

A. 2
B. 3
C. 4
D. 5

MCQ / MSQ

upGrad

How many layers Deep learning algorithms are constructed?

A. 2
B. 3
C. 4
D. 5

MCQ / MSQ

upGrad

The number of nodes in the input layer is 10 and the hidden layer is 5. The maximum number of connections from the input layer to the hidden layer are

A. 50
B. less than 50
C. more than 50
D. It is an arbitrary value

MCQ / MSQ

upGrad

The number of nodes in the input layer is 10 and the hidden layer is 5. The maximum number of connections from the input layer to the hidden layer are

A. 50
B. less than 50
C. more than 50
D. It is an arbitrary value

MCQ / MSQ

upGrad

The input image has been converted into a matrix of size 28 X 28 and a kernel/filter of size 7 X 7 with a stride of 1. What will be the size of the convoluted matrix?

A. 20x20
B. 21x21
C. 22x22
D. 25x25

MCQ / MSQ

upGrad

The input image has been converted into a matrix of size 28 X 28 and a kernel/filter of size 7 X 7 with a stride of 1. What will be the size of the convoluted matrix?

A. 20x20
B. 21x21
C. 22x22
D. 25x25

MCQ / MSQ

upGrad

In a simple MLP model with 8 neurons in the input layer, 5 neurons in the hidden layer and 1 neuron in the output layer. What is the size of the weight matrices between hidden output layer and input hidden layer?

A. [1 X 5] , [5 X 8]
B. [5 x 1] , [8 X 5]
C. [8 X 5] , [5 X 1]
D. [8 X 5] , [1 X 5]

MCQ / MSQ

upGrad

In a simple MLP model with 8 neurons in the input layer, 5 neurons in the hidden layer and 1 neuron in the output layer. What is the size of the weight matrices between hidden output layer and input hidden layer?

A. [1 X 5] , [5 X 8]
B. [5 x 1] , [8 X 5]
C. [8 X 5] , [5 X 1]
D. [8 X 5] , [1 X 5]

MCQ / MSQ

upGrad

Which of the following functions can be used as an activation function in the output layer if we wish to predict the probabilities of n classes (p1, p2..pk) such that sum of p over all n equals to 1?

A. Softmax
B. ReLu
C. Sigmoid
D. Tanh

MCQ / MSQ

upGrad

Which of the following functions can be used as an activation function in the output layer if we wish to predict the probabilities of n classes (p1, p2..pk) such that sum of p over all n equals to 1?

A. Softmax
B. ReLu
C. Sigmoid
D. Tanh

Challenges in DL - Vanishing/Exploding Gradient

upGrad

Vanishing –
As the backpropagation algorithm advances downwards(or backward) from the output layer towards the input layer, the gradients often get smaller and smaller and approach zero which eventually leaves the weights of the initial or lower layers nearly unchanged. As a result, the gradient descent never converges to the optimum. This is known as the vanishing gradients problem.

Exploding –
On the contrary, in some cases, the gradients keep on getting larger and larger as the backpropagation algorithm progresses. This, in turn, causes very large weight updates and causes the gradient descent to diverge. This is known as the exploding gradients problem.

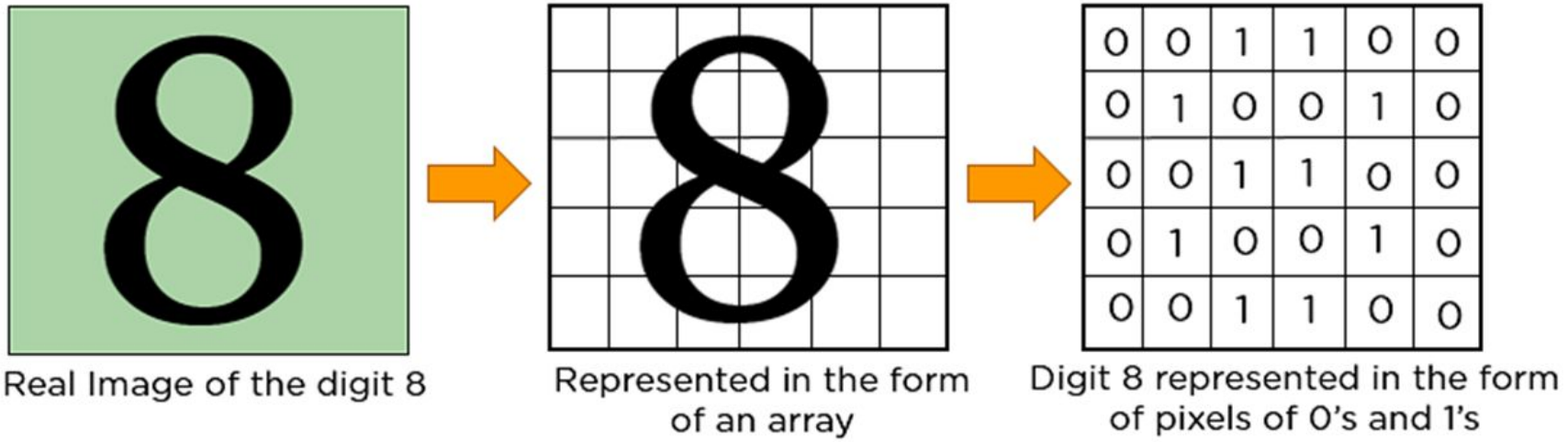
Sample Question

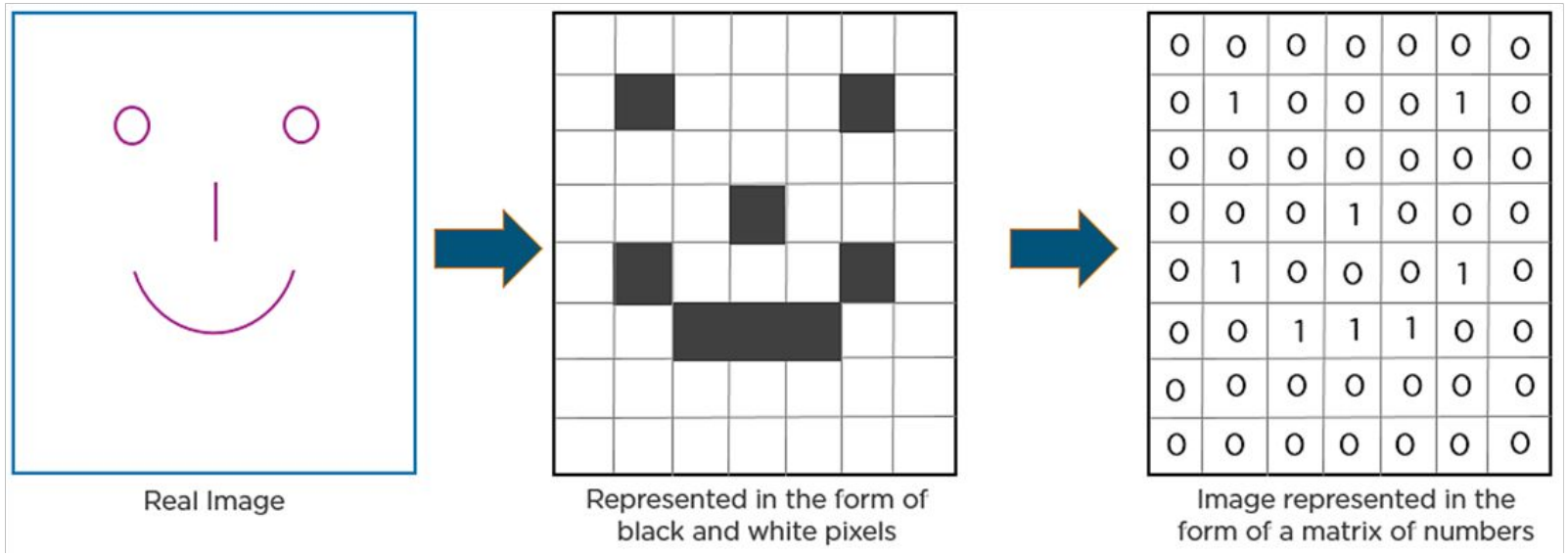
upGrad

- Explain types of neural networks in detail.
- Explain how neural network topology is designed.
- Explain the role of bias in neural networks.
- Explain learning rate in neural networks.
- Explain the role of learning rate in neural networks.
- Explain why weights are modified in Neural Networks.
- Design a Neural Network for diabetes prediction.
- Numeric implementation of Neural Networks (sum).
- Explain Epoch Updating in Neural Networks.
- Differentiate case updating and epoch learning.
- Explain Activation Function in Neural Networks.
- Explain the process of Classification by Backpropagation.
- Explain Backpropagation algorithm in detail.
- Elaborate the possible termination conditions for Neural Network training.
- Explain the challenge of vanishing / exploding gradients.

- Yann LeCun, director of Facebook's AI Research Group, is the pioneer of convolutional neural networks.
- He built the first convolutional neural network called LeNet in 1988. LeNet was used for character recognition tasks like reading zip codes and digits.
- A convolutional neural network is a feed-forward neural network that is generally used to analyze visual images by processing data with grid-like topology. It's also known as a ConvNet.
- A convolutional neural network is used to detect and classify objects in an image.

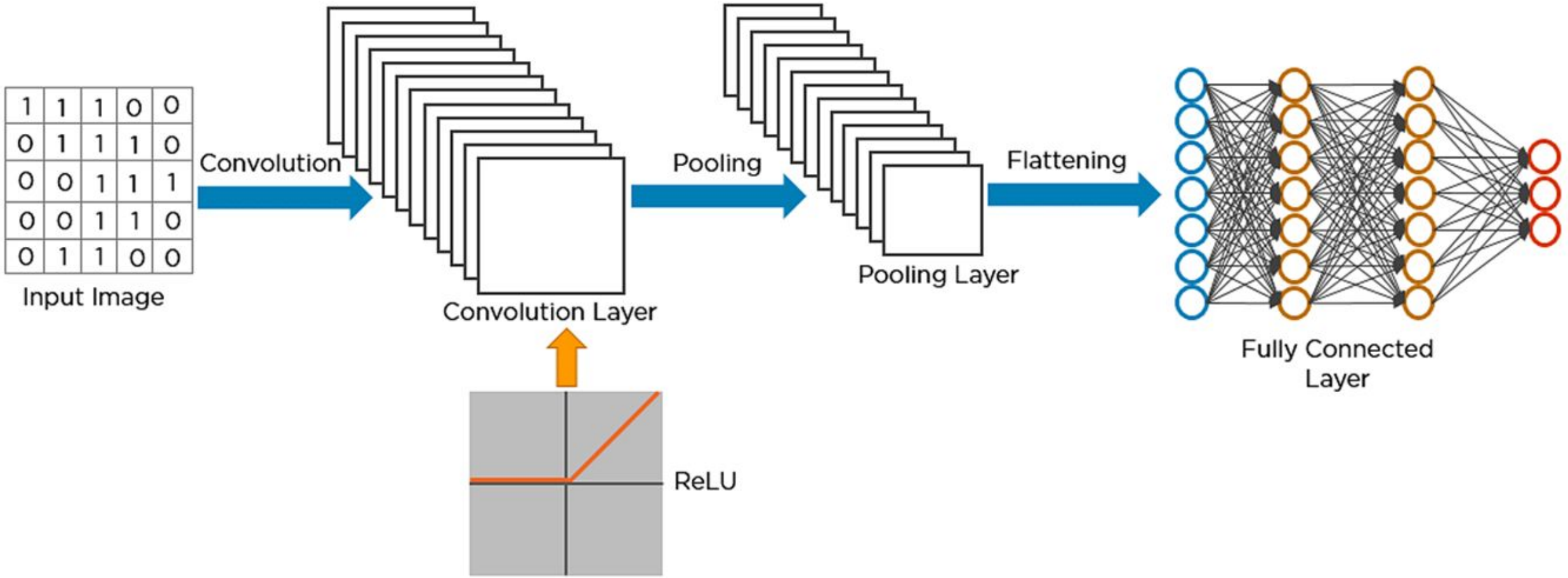
In CNN, every image is represented in the form of an array of pixel values.





CNN - Layers

upGrad



Convolution Operation

upGrad

Input Feature Map

3	5	2	8	1
9	7	5	4	3
2	0	6	1	6
6	3	7	9	2
1	4	9	5	1

Convolutional Filter

1	0	0
1	1	0
0	0	1

Input Feature Map

3×1	5×0	2×0	8	1
9×1	7×1	5×0	4	3
2×0	0×0	6×1	1	6
6	3	7	9	2
1	4	9	5	1

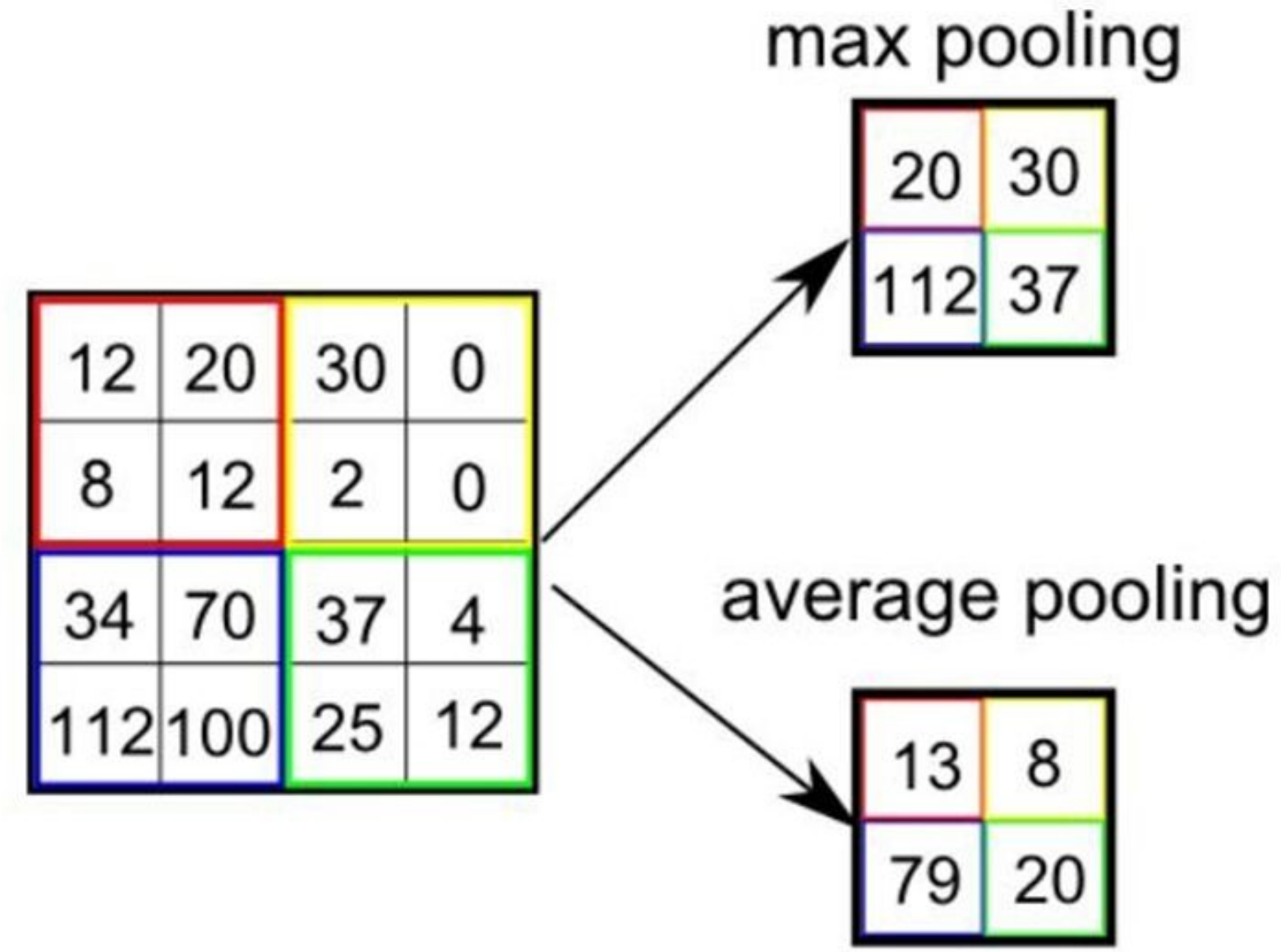
3+0+0+9+7+0+0+0+6

Output Feature Map

25	18	17
18	22	14
20	15	23

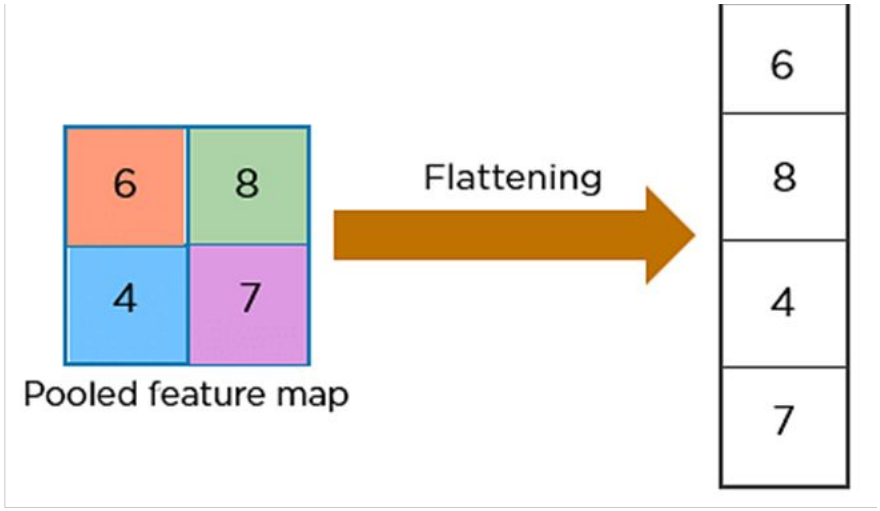
Pooling Layer

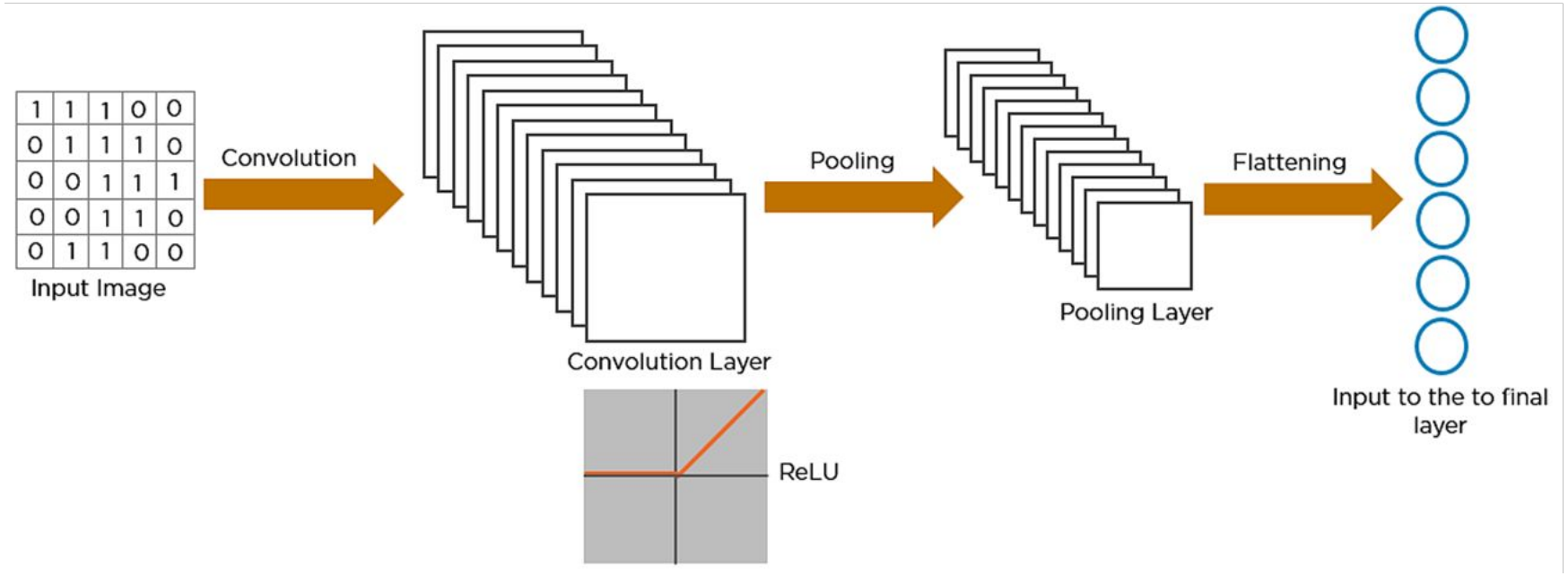
upGrad



Flattening

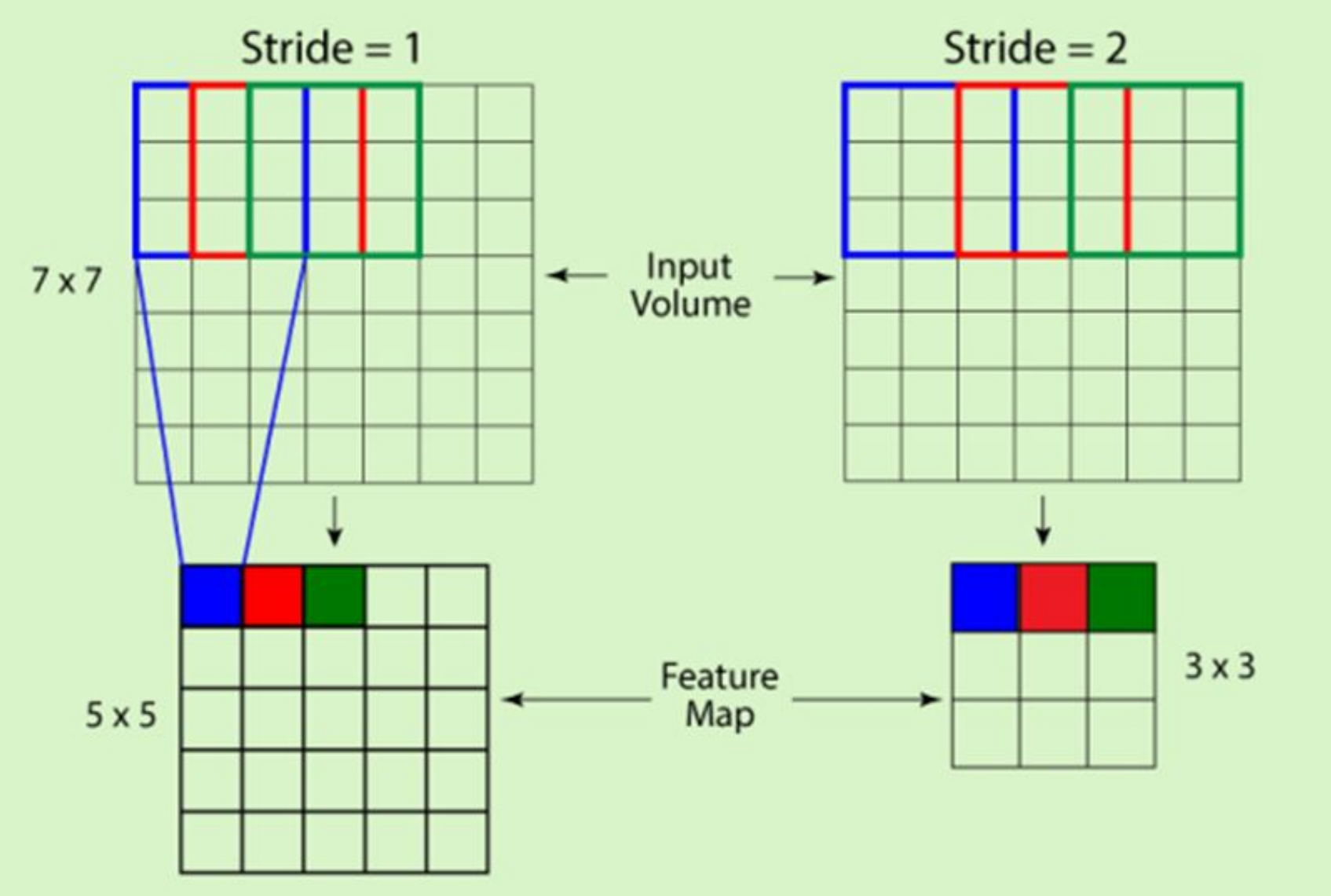
upGrad





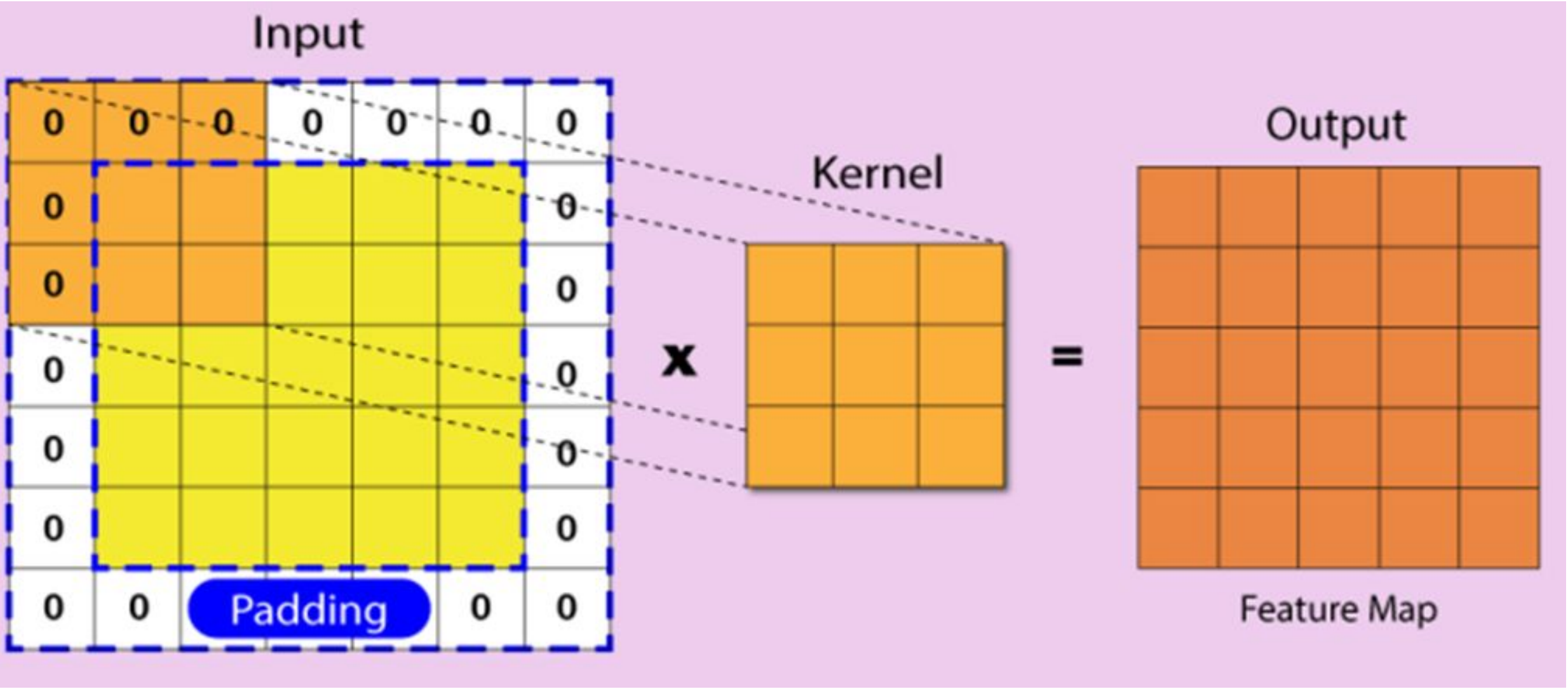
Strides

upGrad



Padding

upGrad



Calculating Output Size

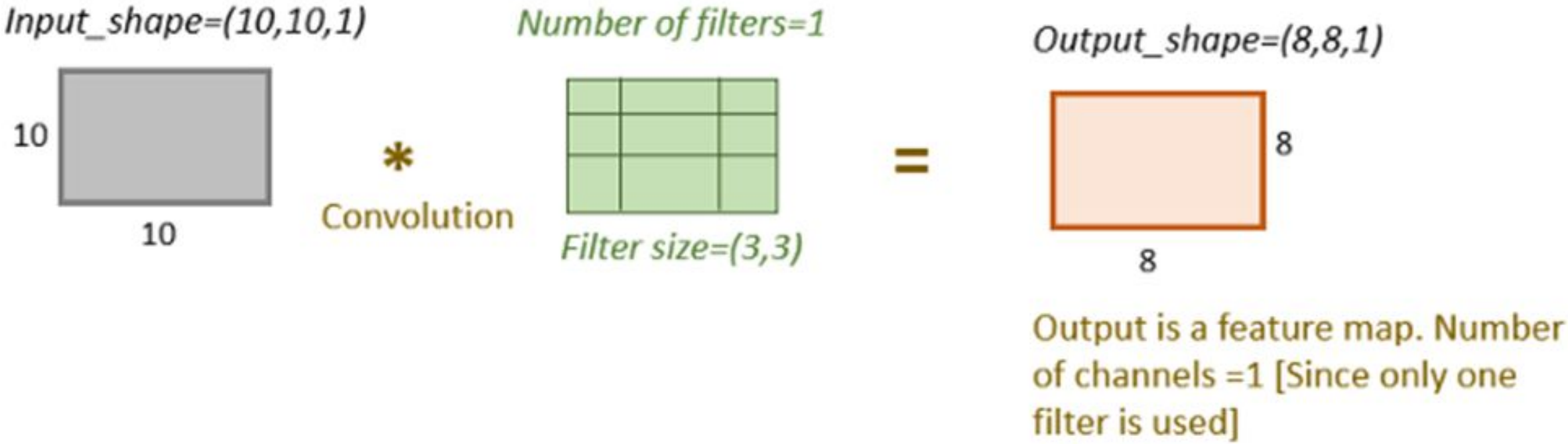
upGrad

$$\text{output size} = \left\lfloor \frac{(\text{input size}) + 2*\text{padding} - (\text{kernel size} - 1) - 1}{\text{stride}} \right\rfloor + 1$$

$$\text{output\_shape} = \frac{n - f + 2p}{s} + 1$$

Calculating Output Size

upGrad



Output is a feature map. Number of channels =1 [Since only one filter is used]

Sample Questions

upGrad

- Explain the issue with going deeper with CNN.
- Explain various layers of CNN.
- Explain the role of Kernels in CNN.
- Explain the need of Padding in CNN.
- Explain types of Padding in CNN.
- Apply Max Pooling and Average Pooling on the given Feature Map using Pool Size 2 x 2.
- Explain Strides in context of CNN.
- Explain Flattening in CNN.
- Explain the role of multiple features.
- Apply Convolution operation on a given Input Feature Map using the given Convolution Filter and derive the Output Feature Map (Numerical).
- Explain ReLu activation function.

Sample Questions

upGrad

- Explain the generic architecture of CNN.
- List various CNN architectures.
- Explain hyper-parameters of CNN.
- Explain the role of Inception Module.
- Apply CNN for hand-written digit recognition (digits 0-9) / character recognition (alphabets a-z) & explain its various layers.
- Consider an image of size $12 \times 12 \times 420$ and apply 5×5 convolution having 36 filters — with and without using 1×1 convolution as intermediate.
- Elaborate the benefit and need of inception module.
- Explain Residual Networks.
- Explain the problem with Sigmoid for Multi-class Classification Problems.
- Explain SoftMax activation function.

THANK YOU