ARTIFICIAL INTELLIGENCE AND DEEP LEARNING
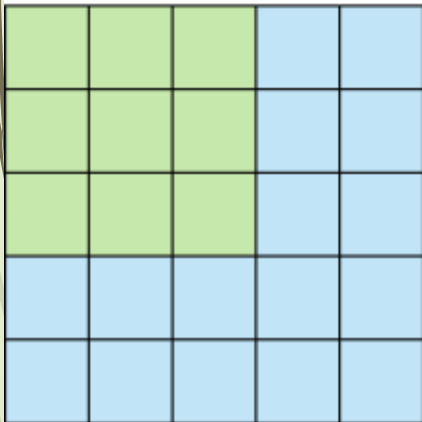
WEEK 8
2021 SPRING
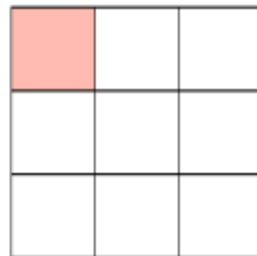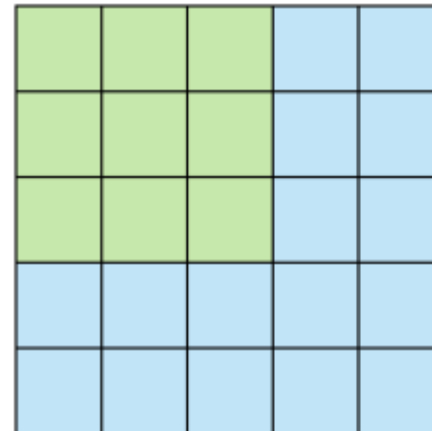
- Stride is the operator that determines the size of the convolution filter at each step. It is generally set to 1 by default. As movement step size increases, the feature map shrinks in size.
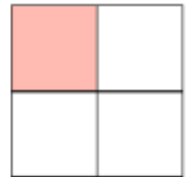

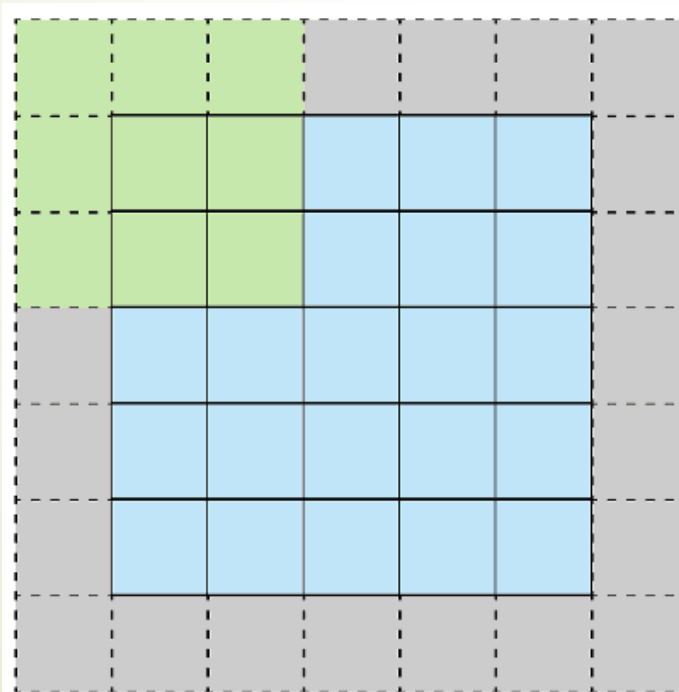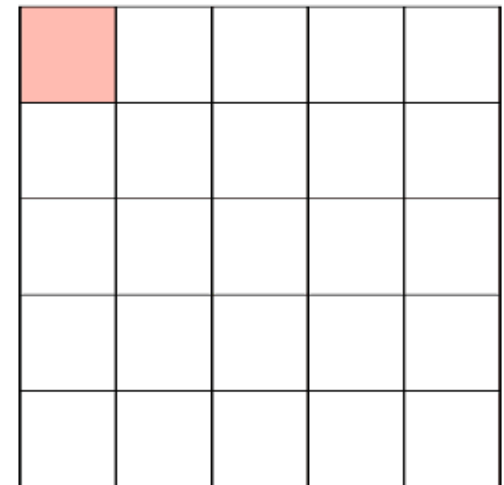Stride=1   Feature Map   Stride=2   Feature Map

- Padding is used to create feature map that is of the same size as input. In this process, cells with value '0' are added around input.


Stride=1 With Padding

Feature Map

- Pooling is applied after the convolution process. With pooling, the size of the feature map is reduced.

- Pooling layer performs sampling by decreasing the width and height of the feature map. It does not change the depth.

- In this context, the most widely preferred method is max pooling. In this method, the window size and stride values are determined.
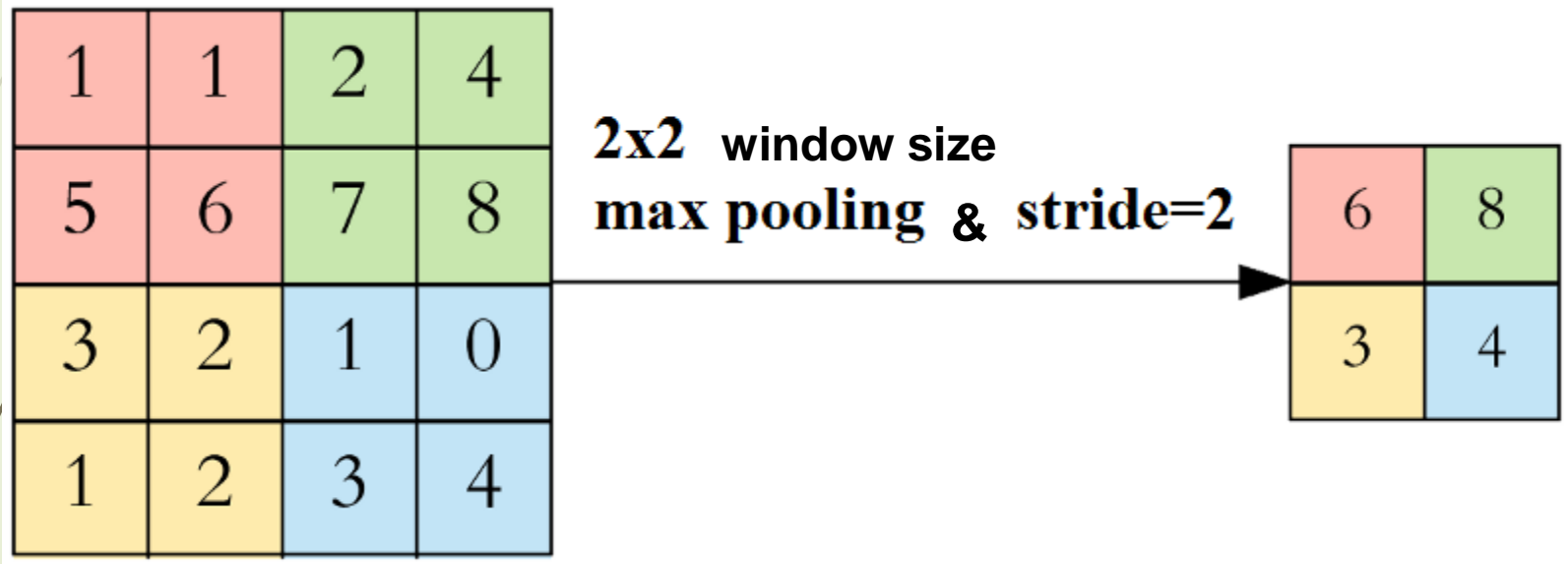
2x2 window size
max pooling & **stride=2**

- In Pooling operation half of the window size and stride feature map is taken. By this way, at the end of pooling process, the width and height of feature map is reduced to half.
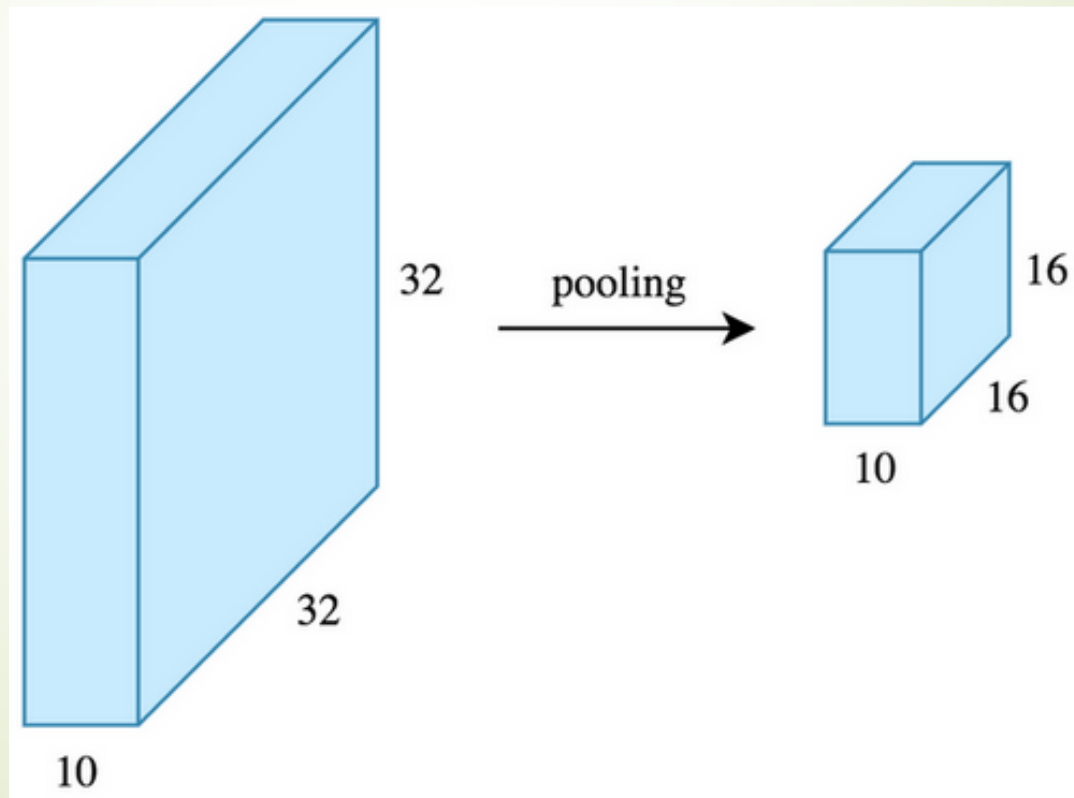
- Fully connected layer follows the pooling layer. The 3 dimensional feature map obtained as the result of pooling layer is converted to a single dimensional vector in this layer. This layer is known as the one where classification is carried out.

- Training of convolutional neural networks is realized by gradient descent method and back propagation.

- In this topic, please visit the web adress below and analyze gradient descent method and back propagation.

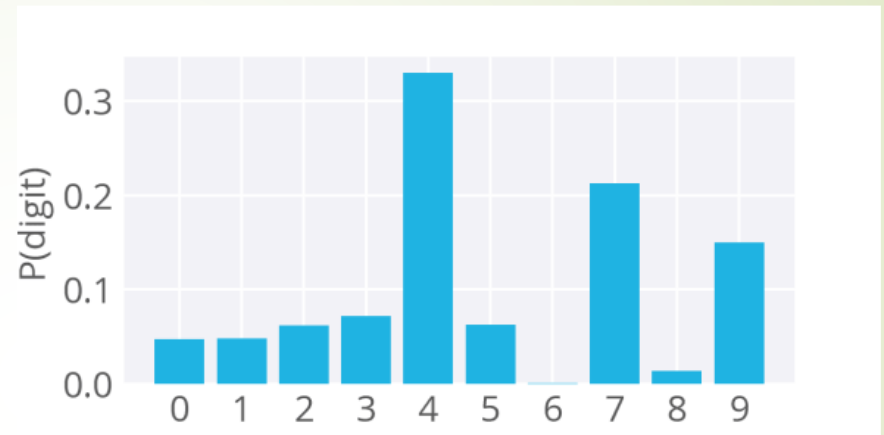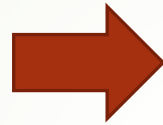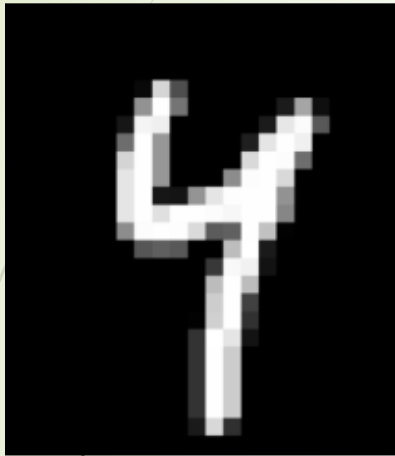http://huseyinatasoy.com/Geri-Yayilimli-Yapay-Sinir-Aglari

- Softmax layer and function are used in classification problems. This function provides the probability of an output to belong to classes.

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}}$$

$$\begin{bmatrix} 1.2 \\ 0.9 \\ 0.4 \end{bmatrix} \longrightarrow \text{Softmax} \longrightarrow \begin{bmatrix} 0.46 \\ 0.34 \\ 0.20 \end{bmatrix}$$

- In convolutional neural networks there are parameters that are not learned directly but determined during the formation of network structure.

- Filter size, number of filters, stride, padding (used or not) are examples of such parameters.