Consider a comparison between a sigmoid function and a rectified linear unit (ReLU). Which of following statement is NOT true?

- Sigmoid function is more expansive to compute
- ReLU has non-zero gradient everywhere
- Sigmoid has a large zone that has zero gradient
- It is possible to compute the “gradient” of ReLU

Consider a comparison between a sigmoid function and a rectified linear unit (ReLU). Which of following statement is NOT true?

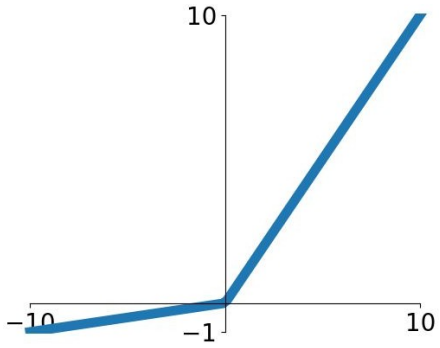
- Sigmoid function is more expansive to compute
- ReLU has non-zero gradient everywhere
- Sigmoid has a large zone that has zero gradient
- It is possible to compute the "gradient" of ReLU

A Leaky ReLU is defined as $f(x)=max(0.1x, x)$. Does it have non-zero gradient everywhere?

- Yes
- No

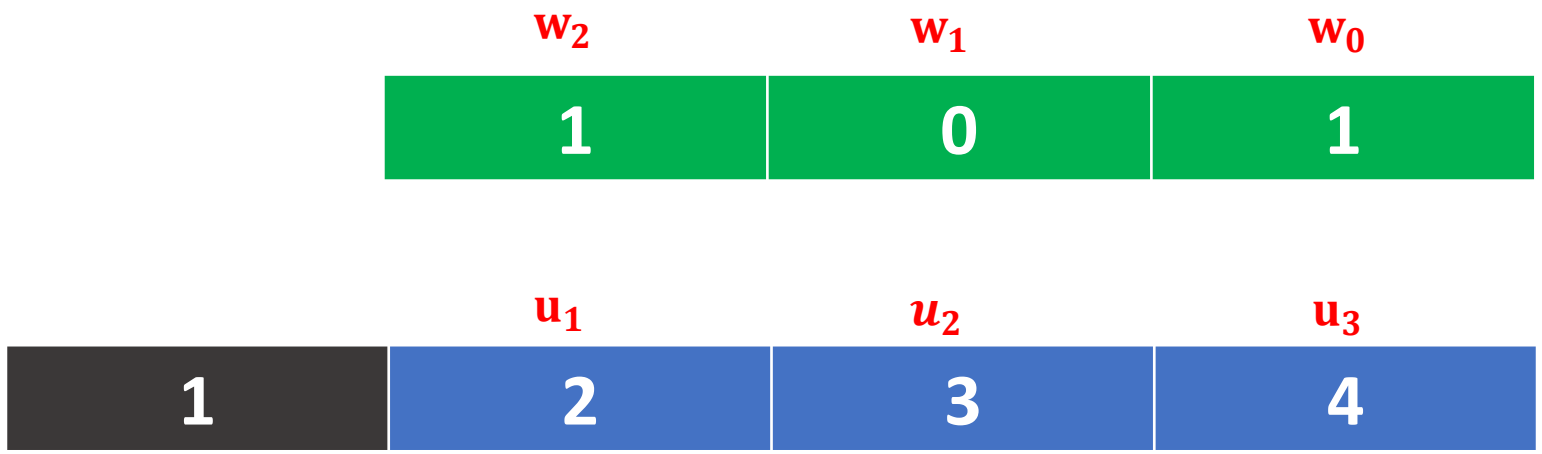
A Leaky ReLU is defined as *f(x)=max(0.1x, x).* Does it have non-zero gradient everywhere?

- Yes
- No



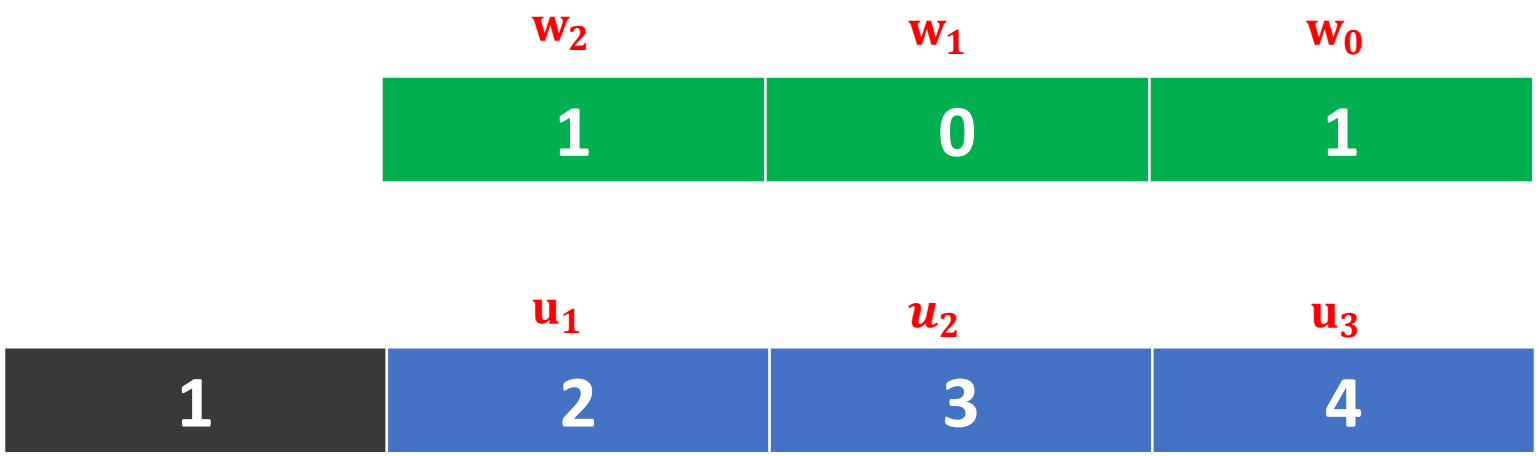
What is the output value of the given convolution operation at the current step?

- 6
- 4
- 9
- 2



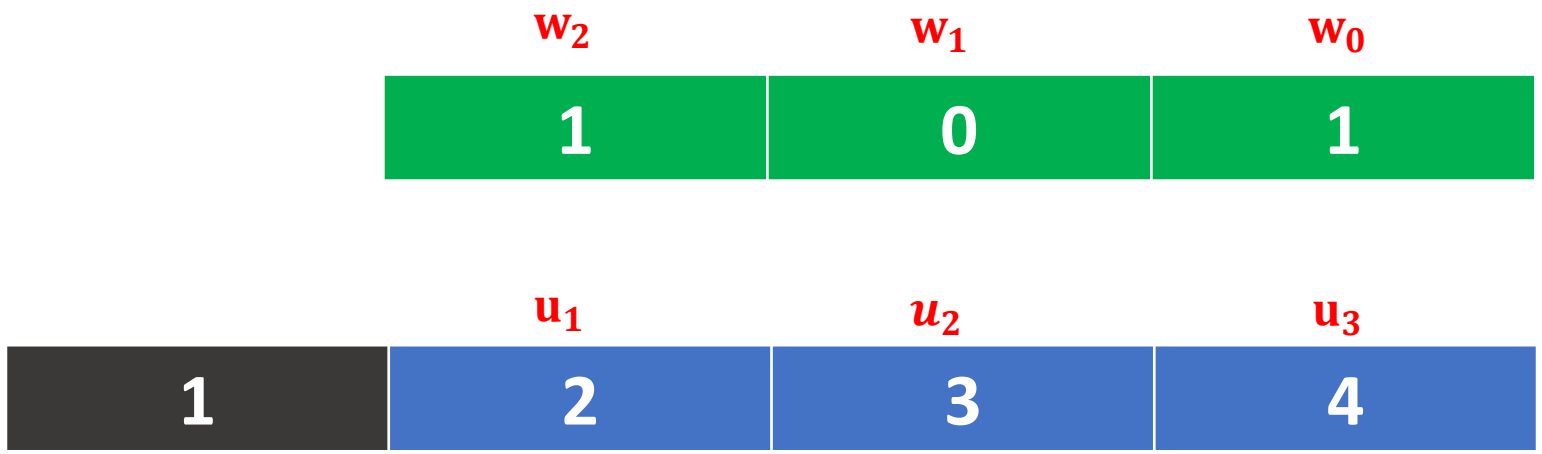
What is the output value of the given convolution operation at the current step?

- 6
- 4
- 9
- 2



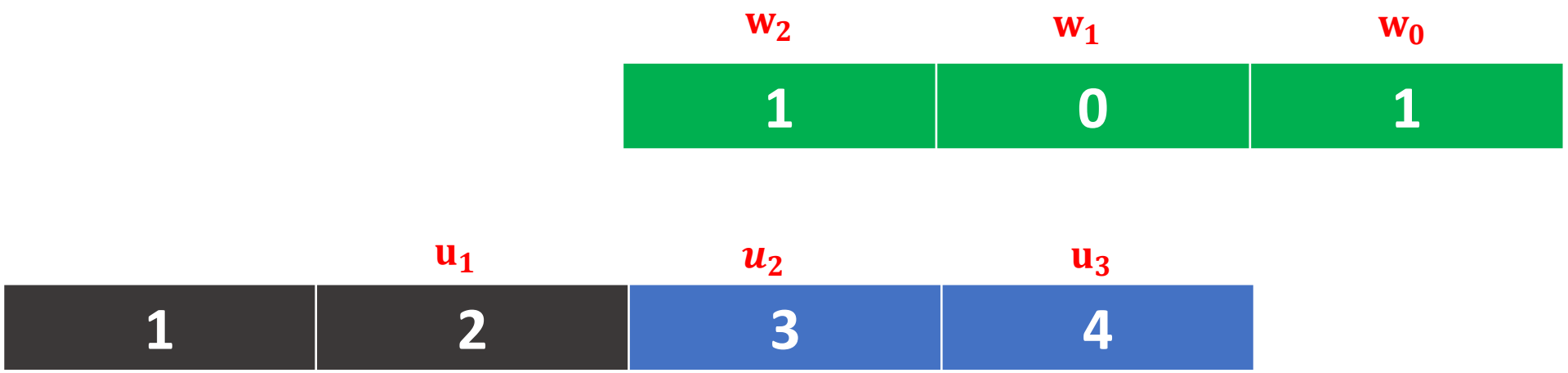
What is the output value of the given convolution operation at the **next** step?

- 5
- 4
- 3
- 2



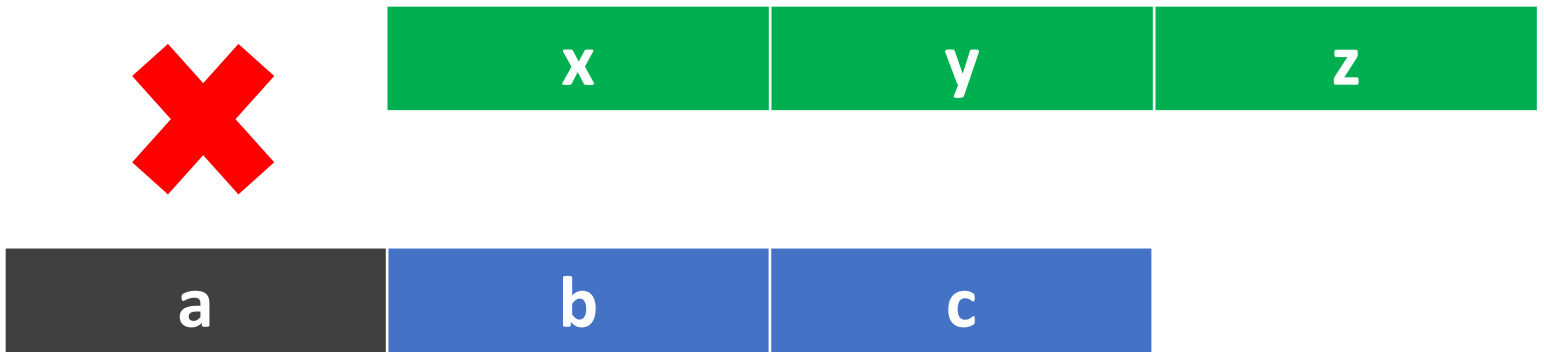
What is the output value of the given convolution operation at the next step?

- 5
- 4
- 3
- 2



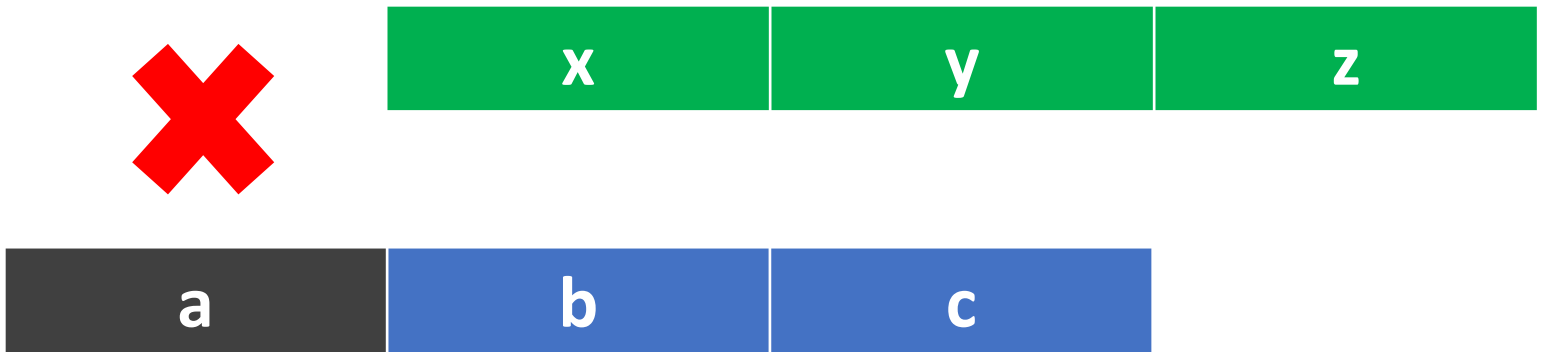
Given an input 1D array of size 7, a convolutional kernel of size 3 with stride 1. If we don’t allow the kernel to partly fall outside of the input, what is the output size?

- 5
- 4
- 7
- 6



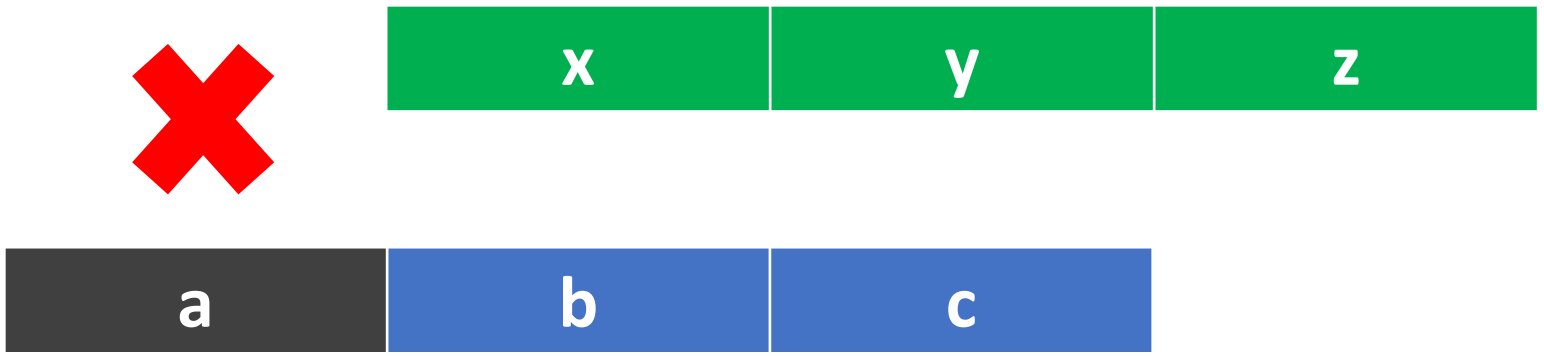
Given an input 1D array of size 7, a convolutional kernel of size 3 with stride 1. If we don’t allow the kernel to partly fall outside of the input, what is the output size?

- 5
- 4
- 7
- 6



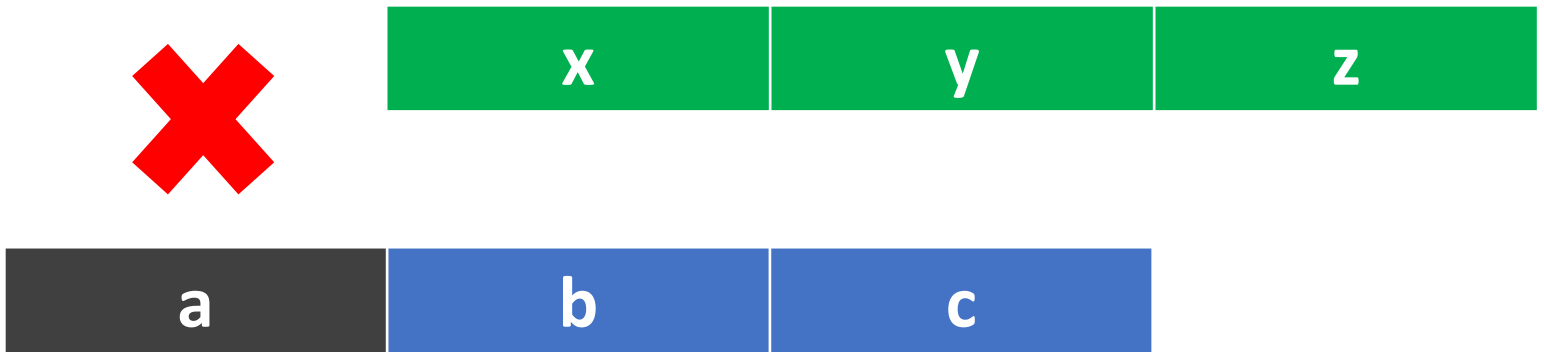
Given an input 1D array of size 7, a convolutional kernel of size 3 with stride **2.** If we don't allow the kernel to partly fall outside of the input, what is the output size?

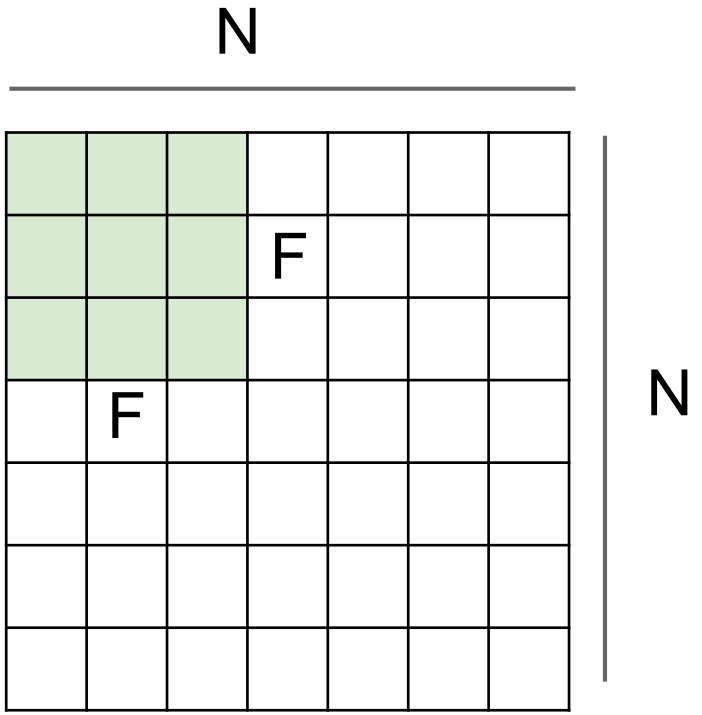
- 6
- 5
- 4
- 3



Given an input 1D array of size 7, a convolutional kernel of size 3 with stride 2. If we don’t allow the kernel to partly fall outside of the input, what is the output size?

- 6
- 5
- 4
- 3

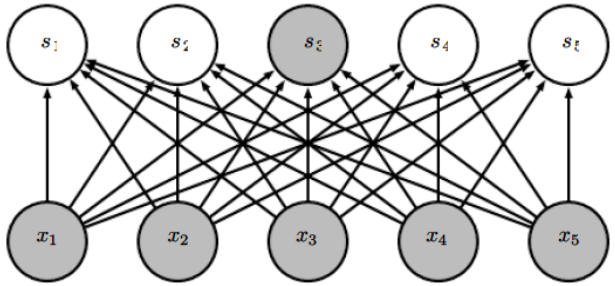




Valid Output size:
(N - F) // stride + 1

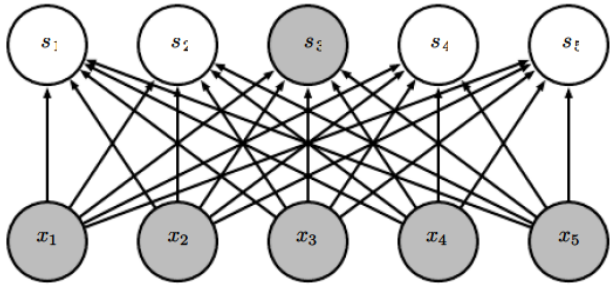
e.g. N = 7, F = 3:
stride 1 => (7 - 3)//1 + 1 = 5
stride 2 => (7 - 3)//2 + 1 = 3
stride 3 => (7 - 3)//3 + 1 = 2

Let us compare a convolutional layer vs. a standard fully connected layer. Which of the following is TRUE?



- Convolution layer has more parameters
- Fully connected layer can be used to represent the convolution
- Convolution layer can be used to represent fully connected layer
- Fully connected layer is more efficient

Let us compare a convolutional layer vs. a standard fully connected layer. Which of the following is TRUE?



- Convolution layer has more parameters
- Fully connected layer can be used to represent the convolution
- Convolution layer can be used to represent fully connected layer
- Fully connected layer is more efficient