Rectified Linear Unit (ReLU)

The rectified linear unit or ReLU activation function is illustrated in Figure 29-9 and works by setting the activation to 0 for values, x, less than 0 and a linear slope of 1 when values, x, are greater than 0.

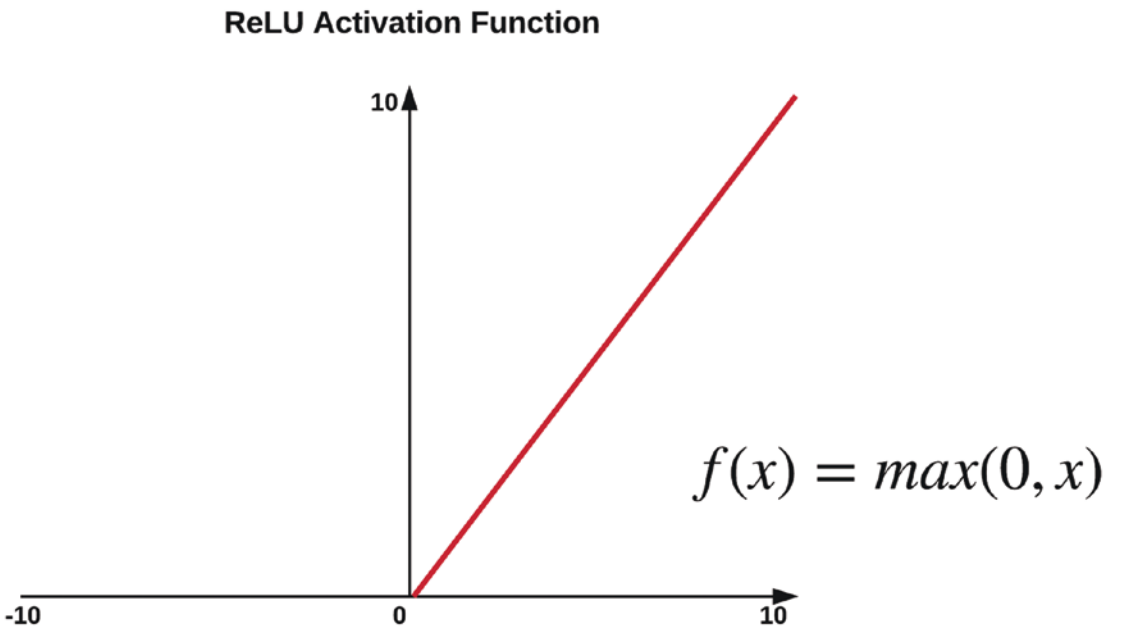

Figure 29-9. *ReLU activation function*

ReLU offers a vast improvement on the tanh and sigmoid activation functions by greatly mitigating the vanishing and exploding gradient problem. However, some gradients can still die out during backpropagation with a large learning rate. However, with a well-defined learning rate, we should not have a problem.

Leaky ReLU

Leaky ReLU is another activation function that is proposed to solve the case of some neurons completely dying out in ReLU by avoiding zero gradients. Leaky ReLU is illustrated in Figure 29-10. The function works by setting the activation to a small negative slope when the value $x < 0$.