

# All you need to know about different **ACTIVATION FUNCTIONS**

-letthedataconfess →

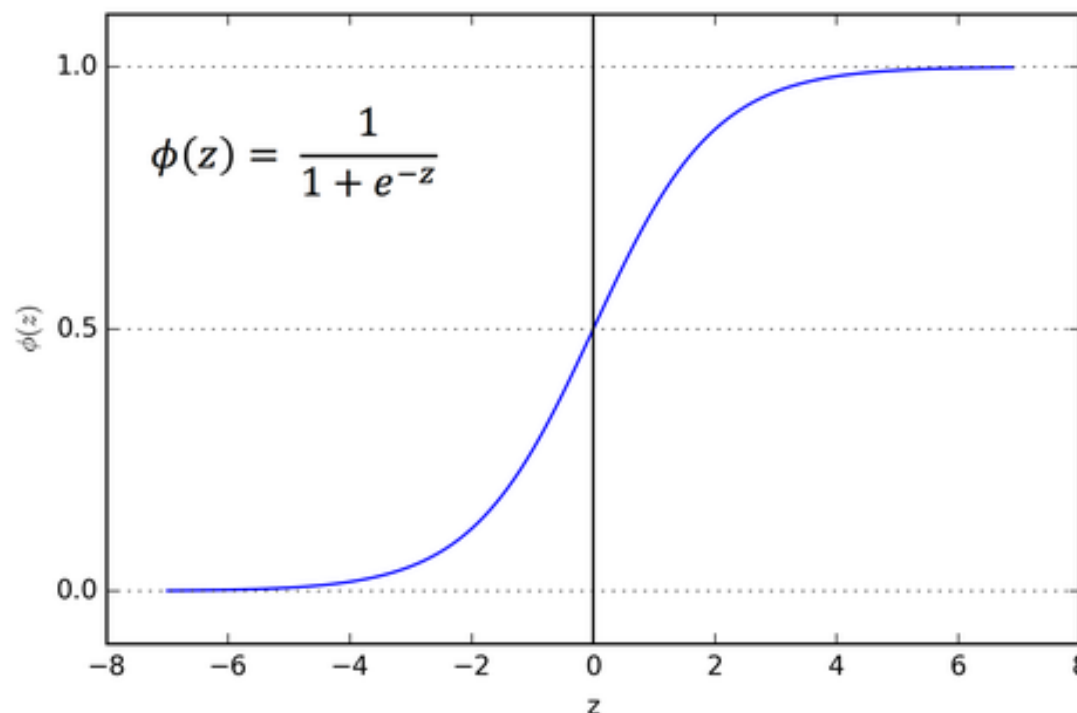
Ace your next data science interview  
Practice Now: <https://www.letthedataconfess.com/mock-interview/>



# Sigmoid function

**Sigmoid** is one of the most popular activation functions used so far. This function looks like an S-shaped curve. The output of the sigmoid will always range between 0 and 1.

This is mainly used in binary classification problems.

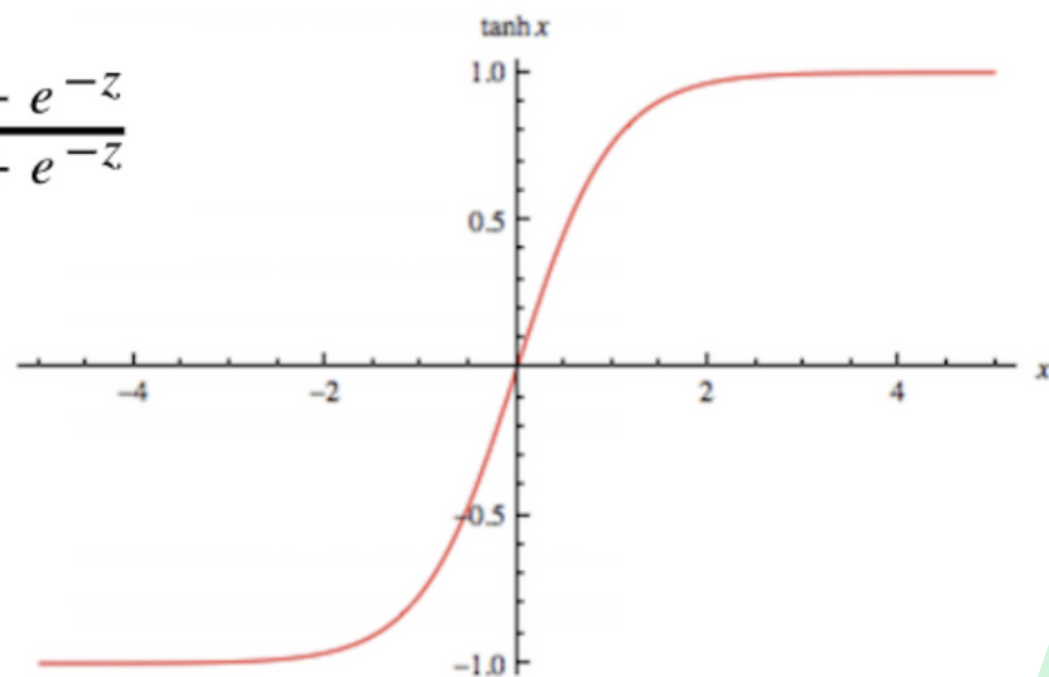


# TanH (Hyperbolic Tangent)

**TanH** also has an S-shaped curve. The output of tanH will range between -1 to 1. As this is a zero centric function, it performs better than sigmoid.

The main advantage of this is that the negative inputs will be mapped strongly to the negative and zero inputs will be mapped to almost zero in the graph of TanH.

$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

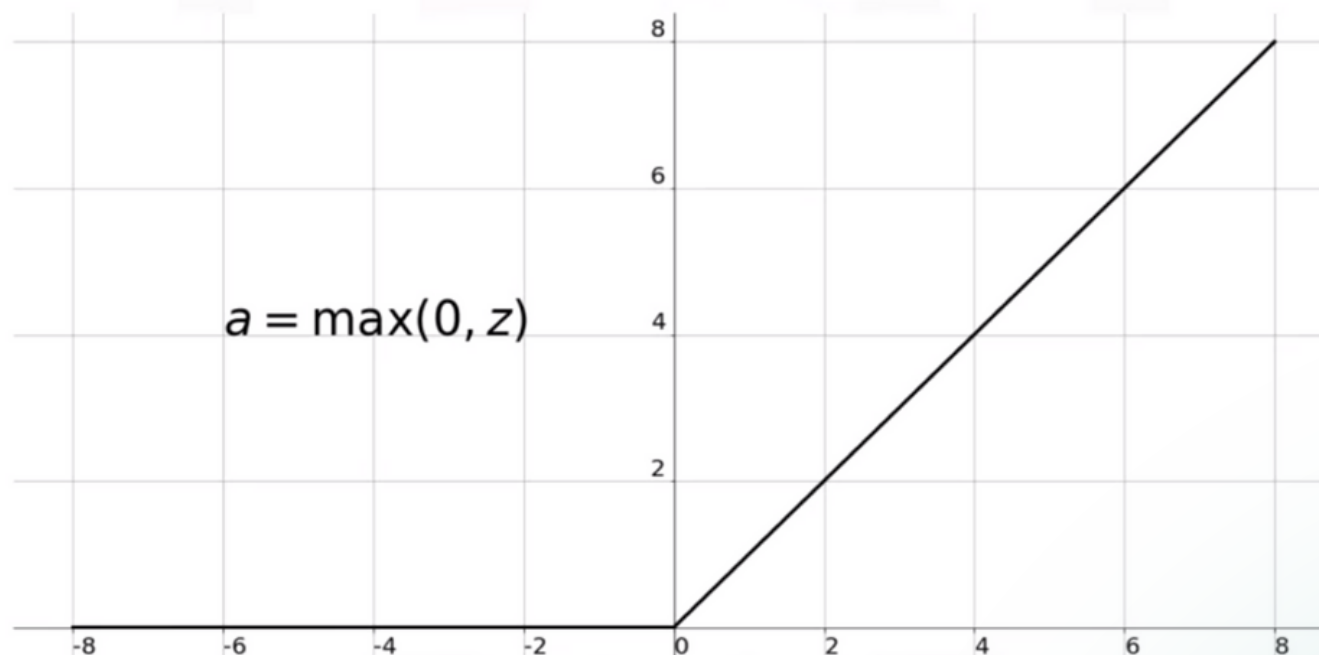


# ReLU (Rectified Linear Unit)

**Relu** is much faster to compute compared to sigmoid and TanH functions.

The Relu only considers the positive values. Negative values are discarded.

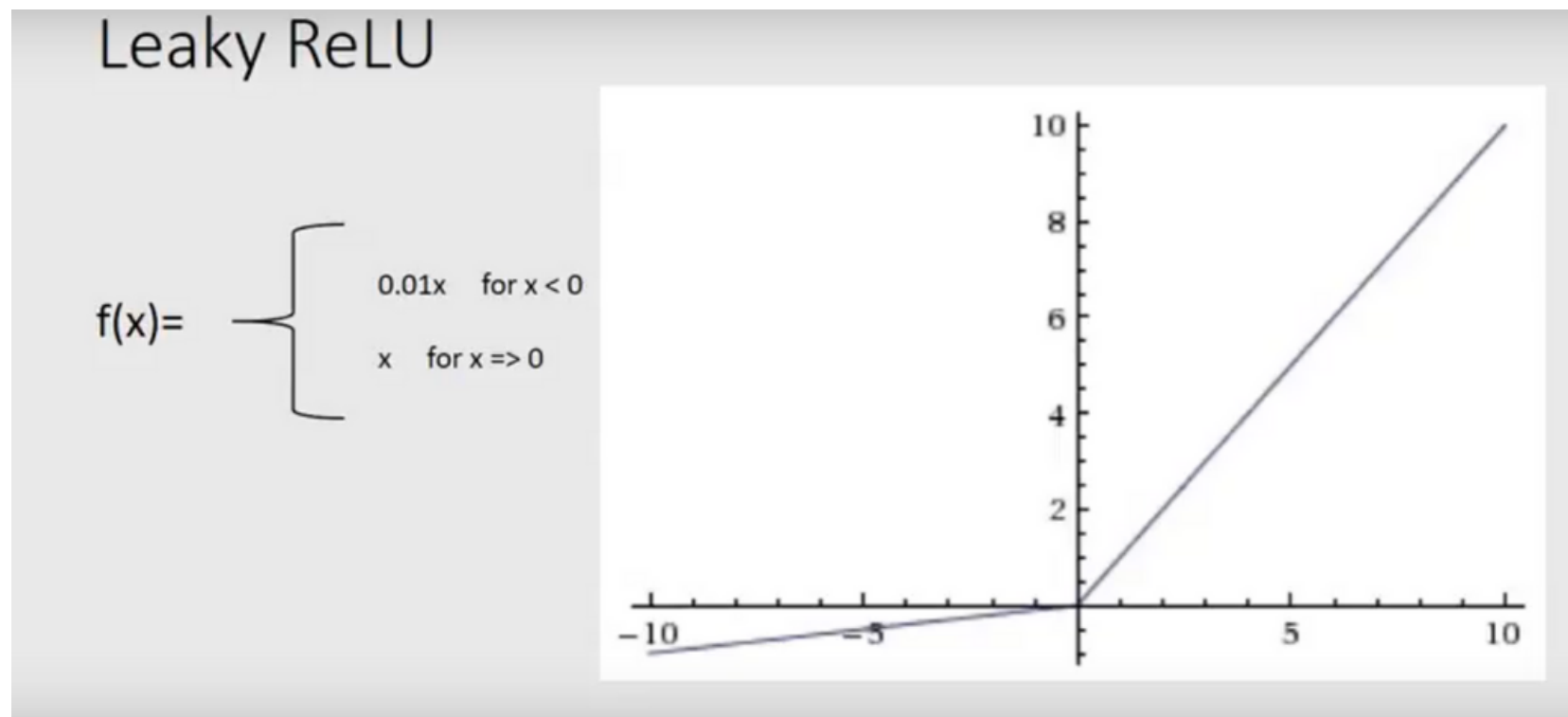
The range of Relu is between 0 to infinity.



# Leaky ReLU

**Leaky Relu** is designed for solving the drawback of Relu which is the dead activation. Leaky Relu also considers a small linear negative value of 0.01

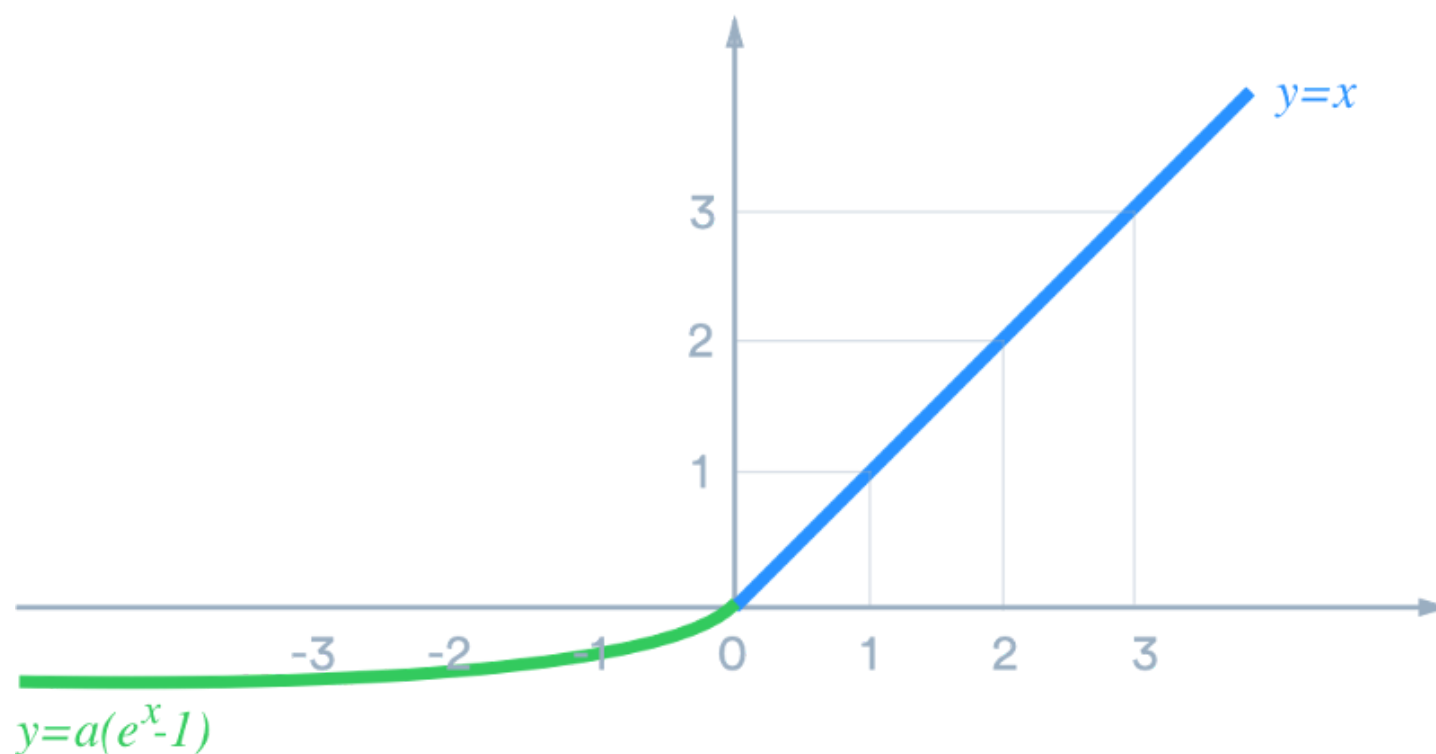
Which helps to increase the range of Relu function from -ve infinity to +ve infinity.



# ELU (Exponential Linear Units)

**Elu** is one of the variations of Relu which also solves the dead Relu problem. Elu, just like leaky Relu also considers negative values by introducing a new alpha parameter and multiplying it with another equation.

Elu is slightly more computationally expensive than leaky Relu.



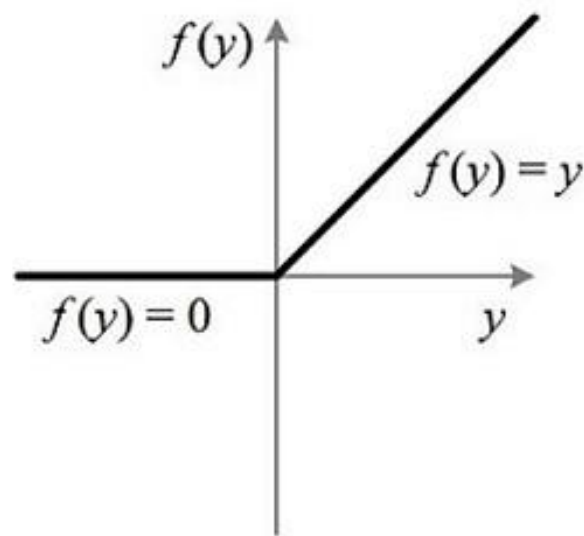
# PRelu (Parametric Relu)

**PRelu** is another variation of Relu. Which improves over Relu. PRelu has a new parameter called 'alpha ' which is a number between 0 and 1

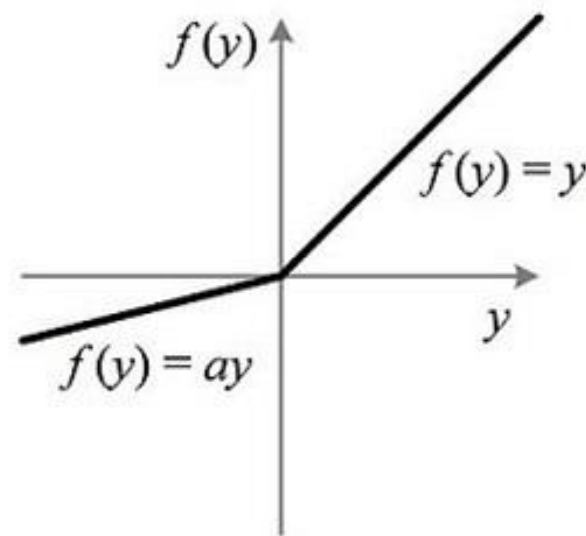
If  $\alpha = 0$ , then PRelu will become Relu

If  $\alpha > 0$ , then PRelu will become leaky Relu.

$\alpha$  is a learnable parameter that is updated during backpropagation.



$$ReLU(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$



$$PReLU(x_i) = \begin{cases} x_i & \text{if } x_i > 0 \\ \alpha_i x_i & \text{if } x_i \leq 0 \end{cases}$$



# Softmax function

**Softmax** is also a well-known activation function. This is mainly used for multi-class classification problems.

The softmax function gives the probability of the current class with respect to others. This means that it also considers the possibility of other classes too.

$$\text{softmax}(z_j) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \text{ for } j = 1, \dots, K$$

