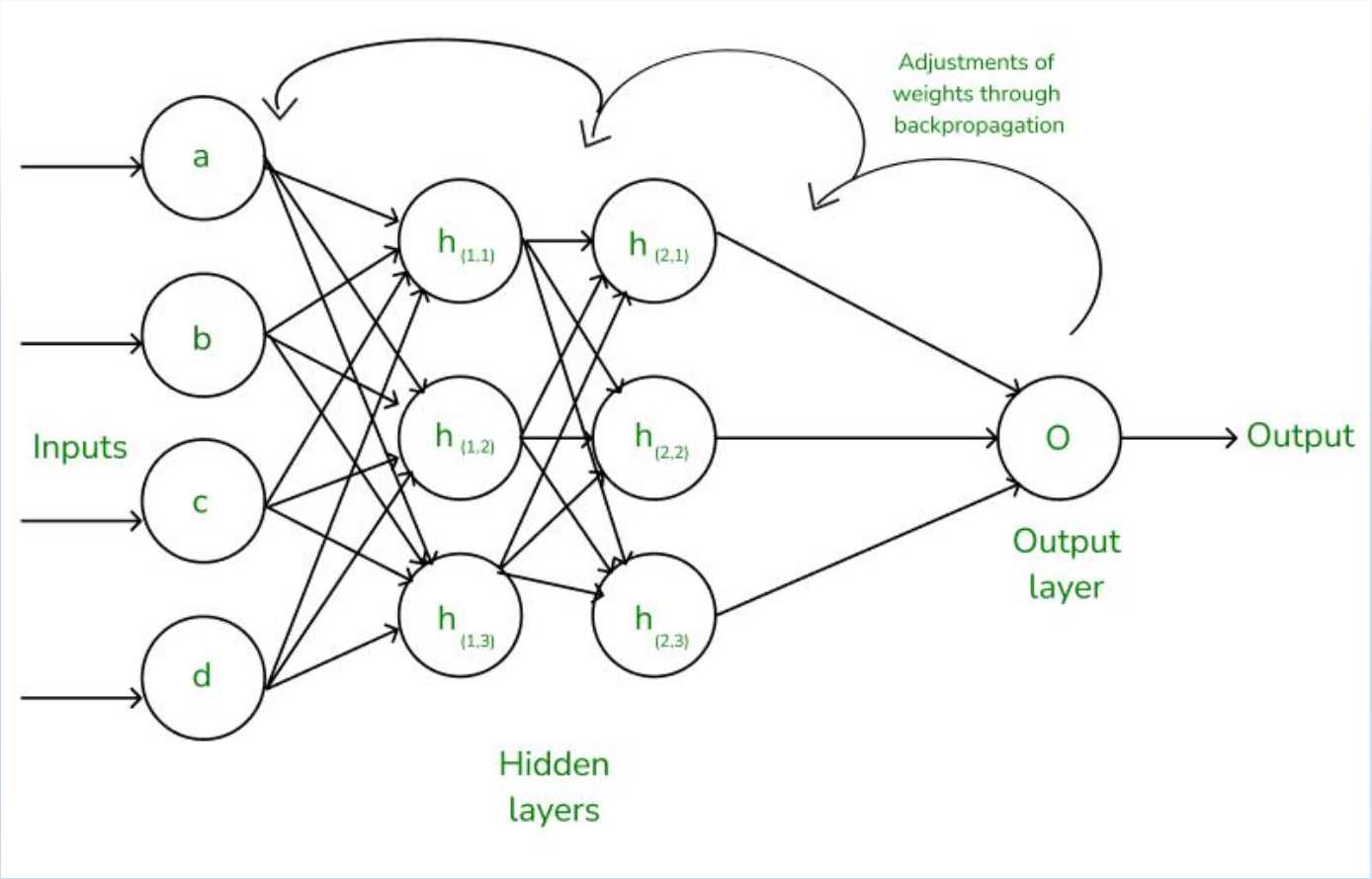
Deep Learning

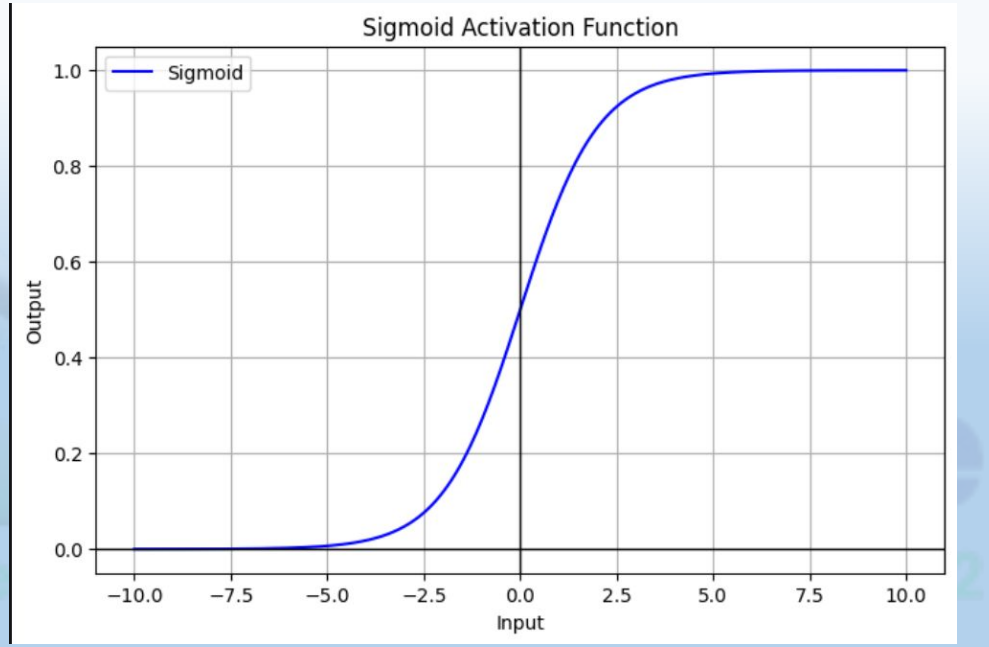
Back propagation

Gananath Bhuyan
Assistant Professor
School of Computer Engineering
KIIT DU, Bhubaneswar

Back Propagation



Sigmoid


Sigmoid Activation Function
Sigmoid
1.0
0.8
0.6
0.4
0.2
0.0
Output
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
Input

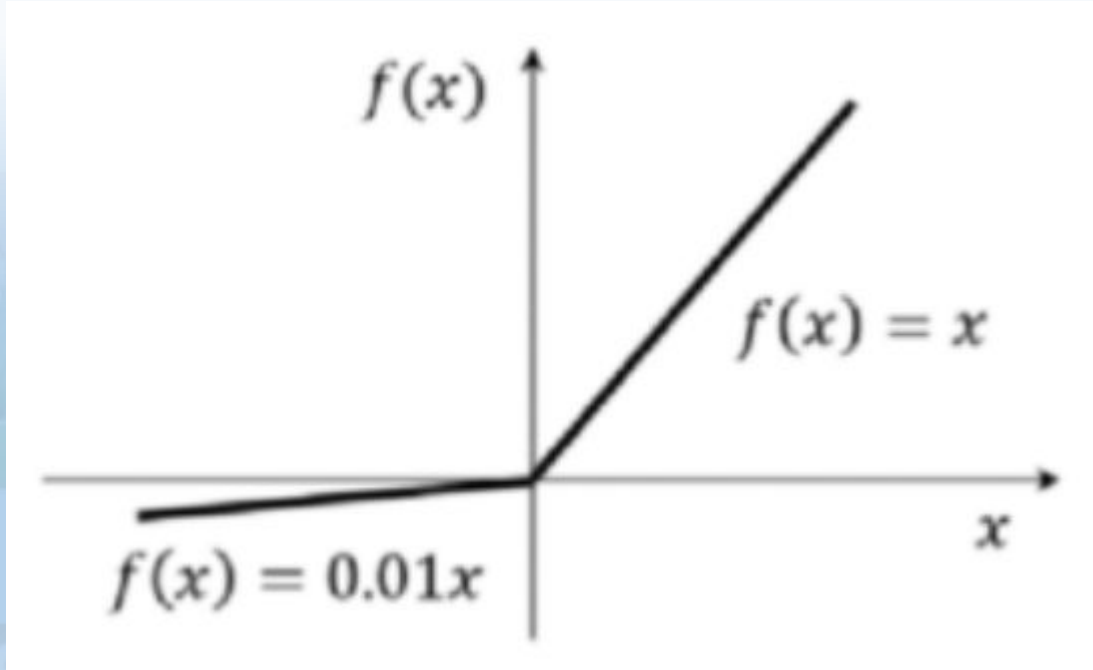
Dying ReLU Problem

- Due to the formula of the ReLU activation function, all the values less than zero will be zero after passing through the ReLU function. If the neural network is having a lot of negative values in the dataset then they all will be converted to zero after passing through the ReLU activation function and there will be zeros in the output which will make the neurons inactive. This type of scenario is called the dying ReLu problem in neural networks

1. Leaky ReLU (LReLU)
2. Parametric ReLU (PReLU)
3. Exponential Linear Unit (ELU)
4. Scaled Exponential Linear Unit (SELU)

Leaky ReLU (LReLU)

if the input passes to the activation function are less than zero then the output value would be 0.01 times the input values and if the input passes to the activation function is greater than zero the output values will be equal to the input value.

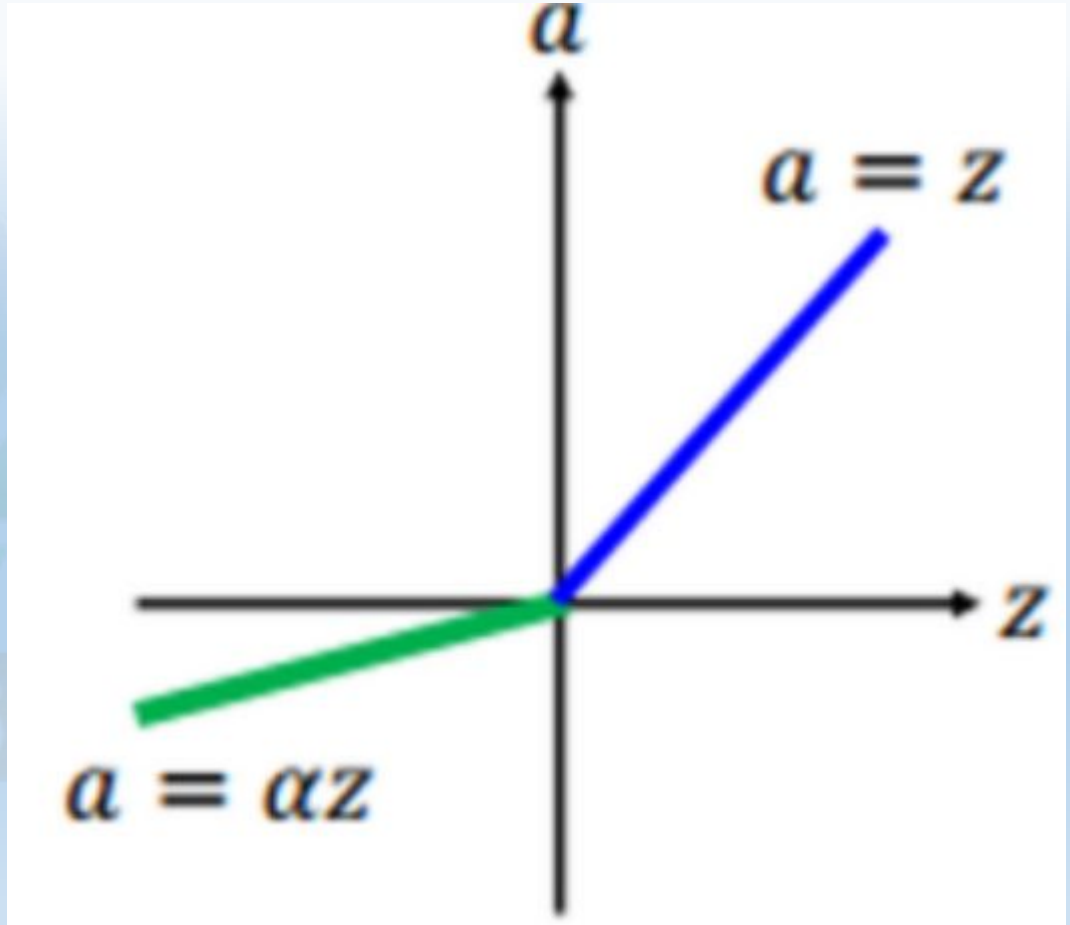


Leaky ReLU(x) = max(0.01x, x)

Parametric ReLU (PReLU)

- if the input value given to the parametric RELU is less than zero then the output values will be equal to the input values multiplied with parameter “a” and if the input value if greater than zero then the output value will be the same as the input value.
- The value of the ‘a’ is learnable and is decided by the neural network itself.

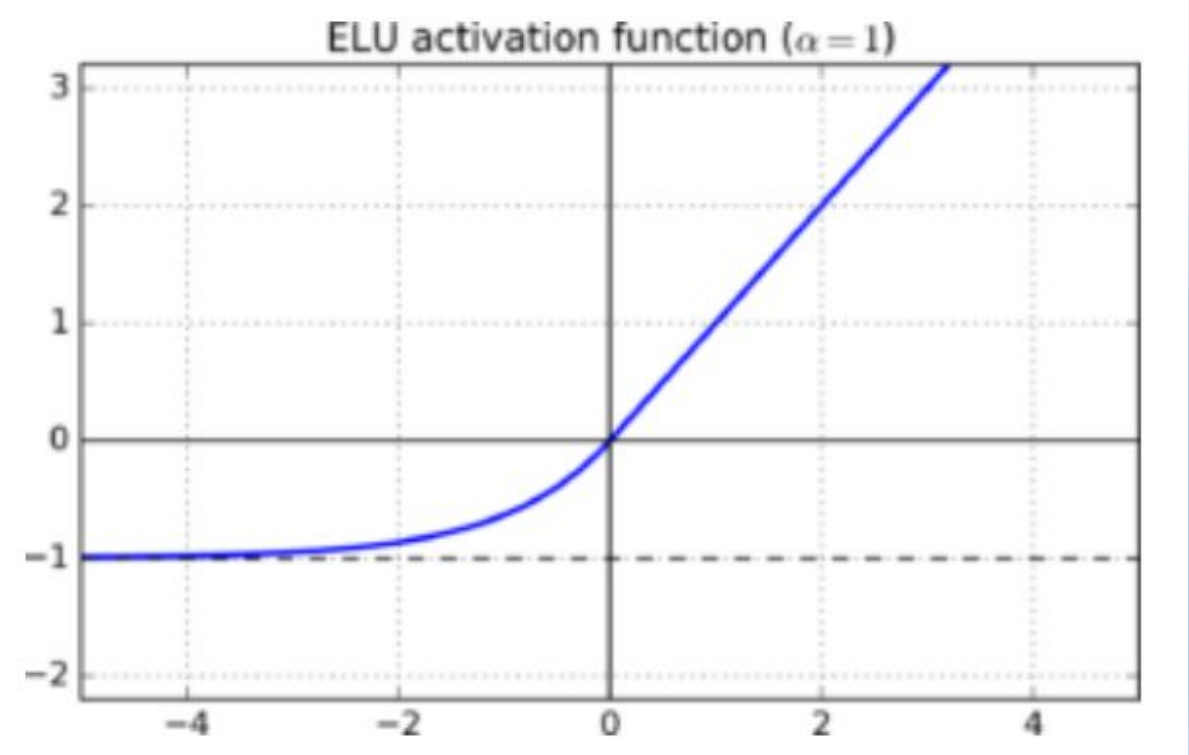
PReLU(x) = max(ax, x)



Exponential Linear Unit (ELU)

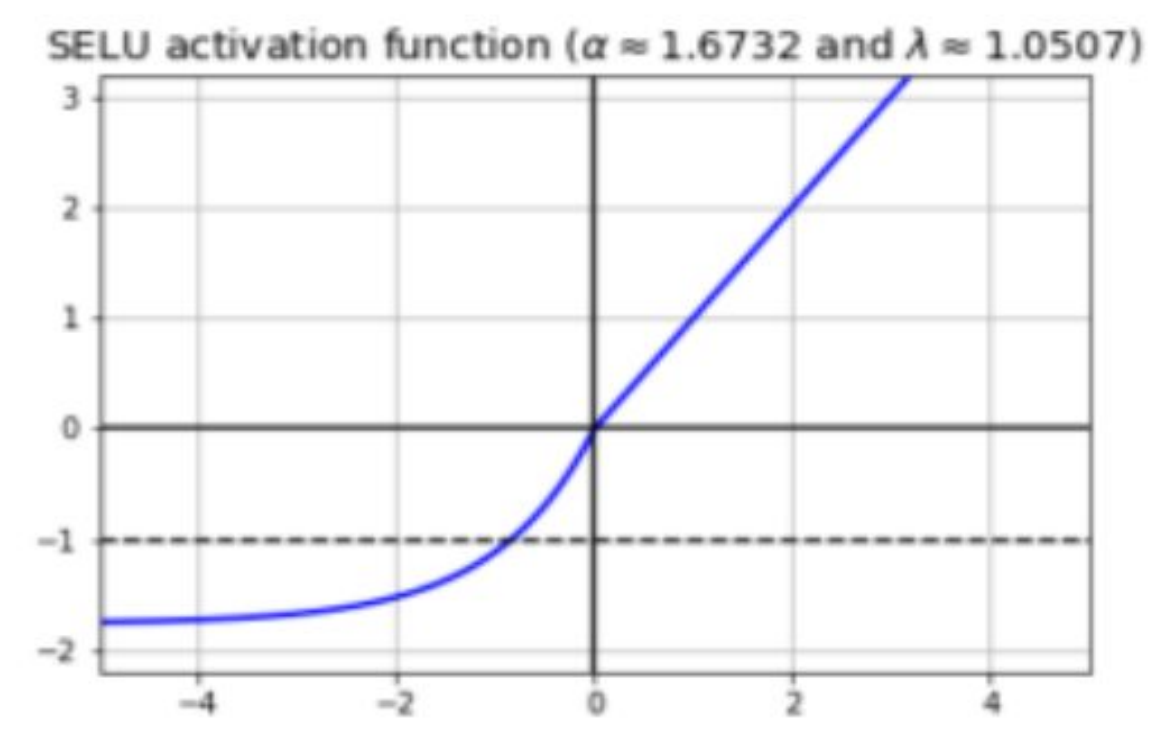
- ELU activation function uses a log curve and due to this it does not gives a straight line for negative values, unlike leaky and parametric ReLU.

$$\begin{cases} x & for\ x \geqslant 0 \\ \alpha\left(e^{x}-1\right) & for\ x < 0 \end{cases}$$



Exponential Linear Unit (ELU)

- if the input value x is greater than zero, the output value becomes x multiplied by lambda λ and If the input value x is less than or equal to zero then alpha is multiplied with the exponential of the x-value minus the alpha, and then we multiply by the lambda value.



$$\text{selu}(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha e^x - \alpha & \text{if } x \leqslant 0 \end{cases}$$

Alpha = 1.6732632423543772848170429916717

Lambda = 1.0507009873554804934193349852946