Name: Advaitesha Gupta

Batch: B2

SAP ID: 500095942

ENROLLMENT NO: R2142211077

COLLEGE: UPES
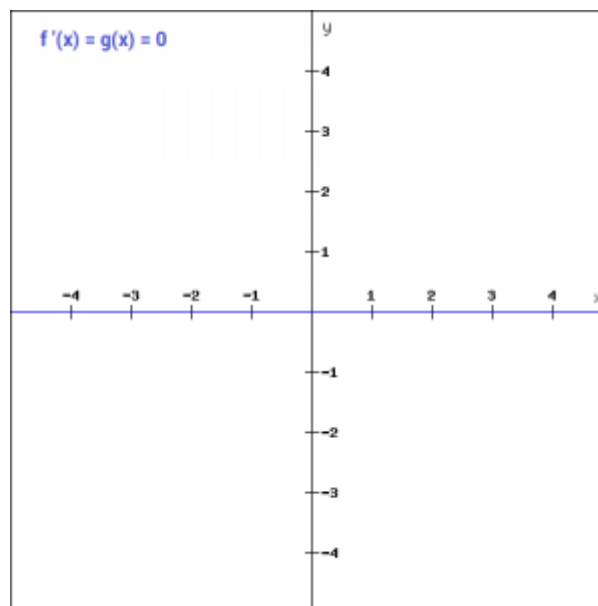
COURSE: B.TECH CSE

BRANCH: AIML-H

ACTIVATION FUNCTIONS

1. **Binary step function:**
- works on threshold value
- used when data is small
- problem: the gradient used for updating of weights and biases is zero

- **Mathematical function**:


2. **Linear function:**
- To avoid problem of zero gradient we keep some component of x axis.
- **Mathematical function**: $f'(x) = a$

- Although the gradient here does not become zero, but it is a constant which does not depend upon the input value x at all. This implies that the weights and biases will be updated during the backpropagation process but the updating factor would be the same.


f(x) = 4x
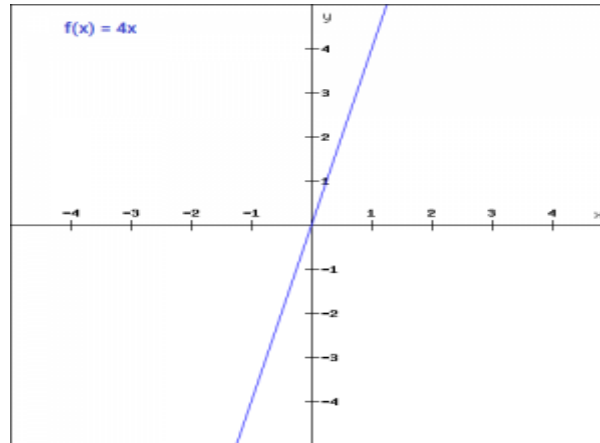
3. Sigmoid function:
- Sigmoid transforms the values between the range 0 and 1.
- Used for non-linear data

- **Mathematical function**: $f(x) = 1/(1+e^{-x})$
- $f'(x) = sigmoid(x)*(1-sigmoid(x))$ where gradient values are significant for range -3 and 3

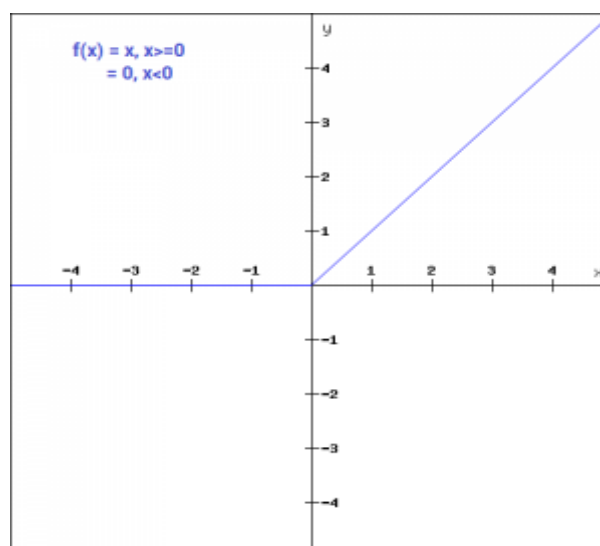f(x)=1/(1+e^-x)
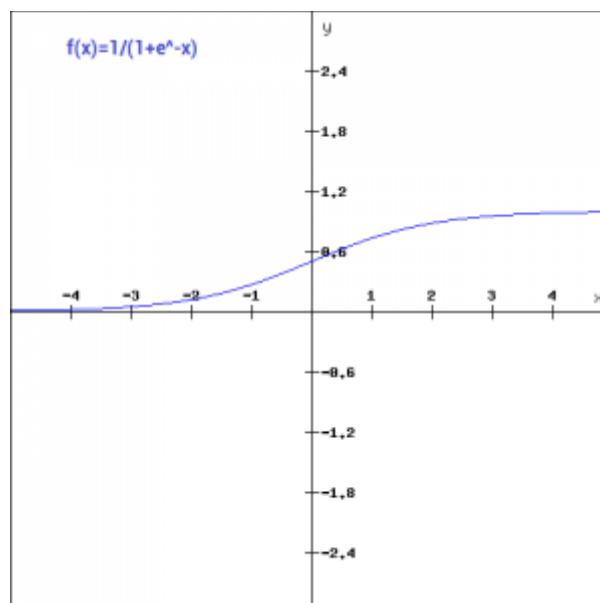
4. **Tanh function:**

- Only difference between sigmoid and tanh is that it is symmetric around the origin. The range of values in this case is from -1 to 1.
- The inputs to the next layers will not always be of the same sign.
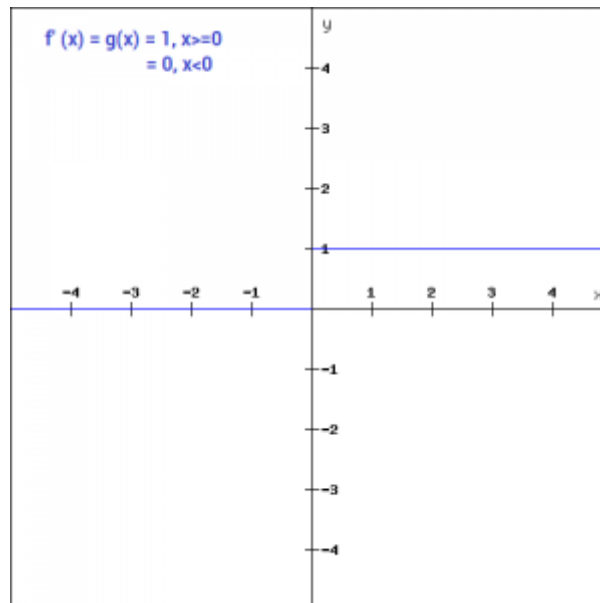- It is generally used by hidden layers.

$$f(x) = 2/(1+e^{(-2x)}) - 1$$
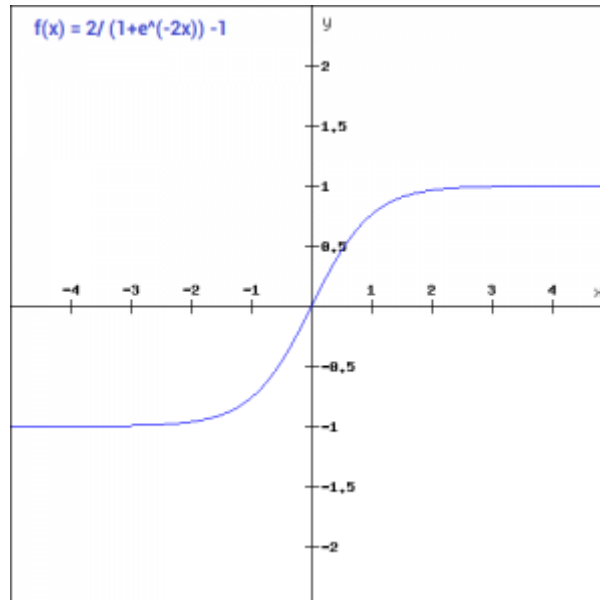
- **Mathematical function**: $\tanh(x)=2\text{sigmoid}(2x)-1$

5. ReLU function:
- The neurons will only be deactivated if the output of the linear transformation is less than 0.
- ReLU stands for Rectified Linear Unit
- It is generally used by hidden layers.

- **Mathematical function:** $f(x)=\max(0,x)$

$$f(x) = x, \; x{\geq}0$$
$$= 0, \; x{<}0$$

Derivative: f'(x) = 1, x>=0
         = 0, x<0