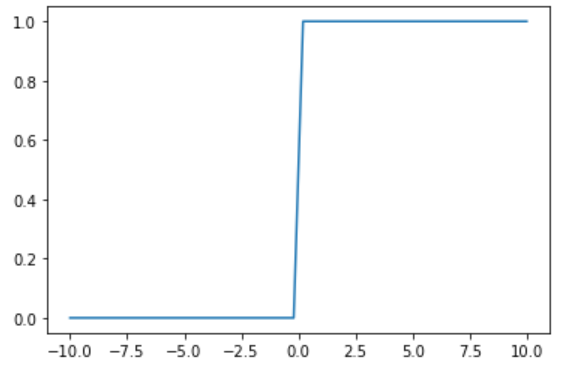
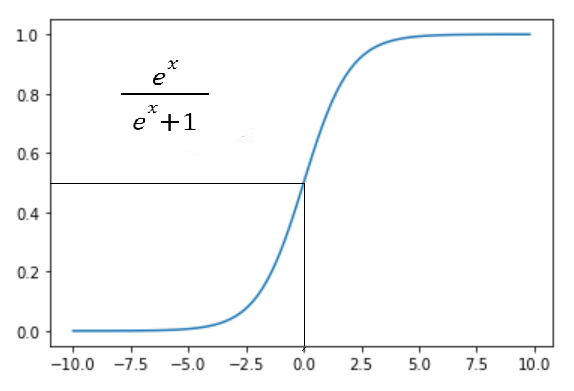
Summarizing Types of Activation Function - Part 1

1. **Step Function**: One of the most basic categories of activation functions is the Step Function. A threshold value is used in this case, and if the net input y value is greater than the threshold, the neuron is activated.



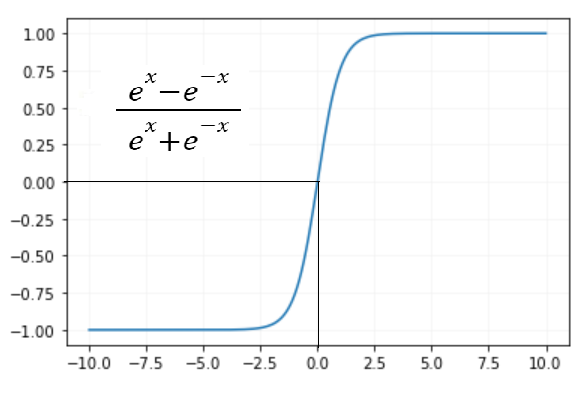
A binary step function is a threshold-based activation function, which means that it activates when a certain threshold is reached and deactivates when it falls below that point. Because there is a sharp jump in the function, the derivative at x=o will give a large number, but that large number is not useful for learning and cannot be used for multi-class classification. Smooth functions are good for learning where there are defined slopes and this step function is a sharp function.

2. **Sigmoid Activation**: This activation function looks like a smoother version of a step function where the slopes at each point are well defined, and this smoothness is caused by using the exponential function.

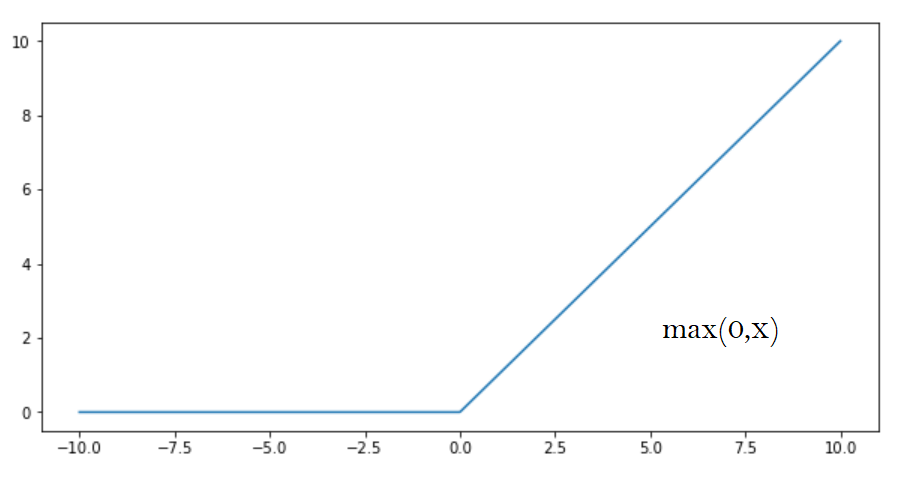


The sigmoid function, also known as the logistic function, has an output that ranges from 0 to 1. It makes use of a probabilistic approach. It is a function that is graphed in the shape of a "S." The sigmoid activation function has a value between 0 and 1. Because the values of the Sigmoid function range between 0 and 1, the outcome can be easily predicted to be 1 if the value is greater than 0.5 and 0 otherwise. In the output layer of a binary classification, Sigmoid function values are typically used, with the result being either 0 or 1.

3. **Tanh Activation**: Tanh activation function is a scaled and compressed version of sigmoid activation function and is also called a hyperbolic tangent. The structure of this function is similar to the sigmoid activation function, but significantly superior because it allows for negative outputs and has an output range of -1 to 1.



4. **ReLU Activation**: Rectified linear activation function is the most commonly used activation function in the hidden layer of a neural network which ranges between 0 to inf.



if X>=0, X

if X<0, 0

If x is positive, it outputs x, and if not, it outputs 0. The ReLU activation function has a range of 0 to inf.

The advantage of ReLU is that it requires fewer mathematical operations than tanh and sigmoid, making it less computationally expensive.

The disadvantage of ReLu is that it produces dead neurons, which never activate, known as the dying ReLu problem.

Happy Learning!!