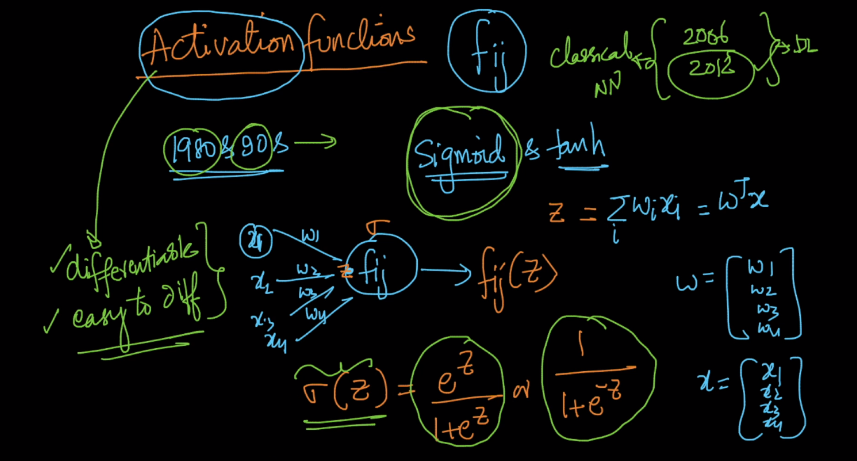
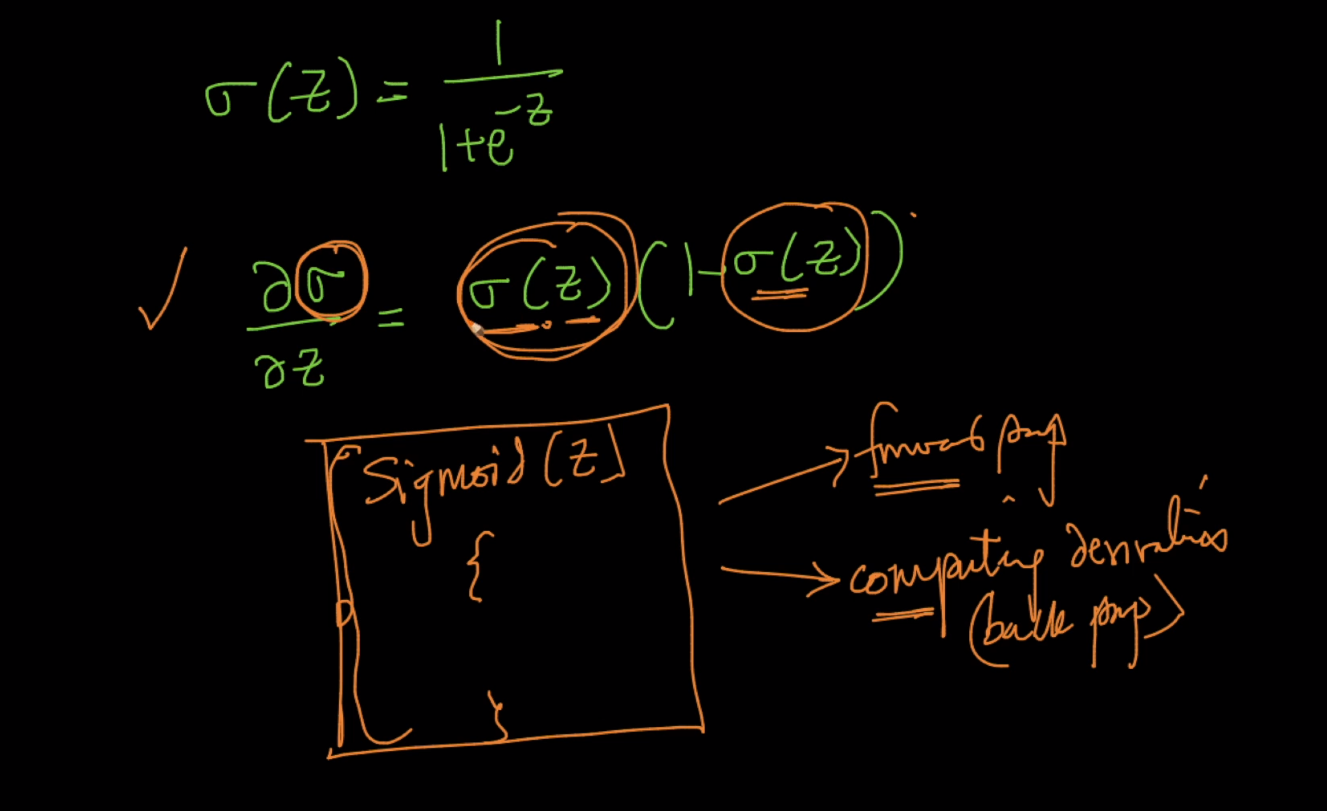
**Activation functions**

Activation function must be differentiable and easy to differentiate, earlier there were two popular activation functions i.e sigmoid and tanh functions.

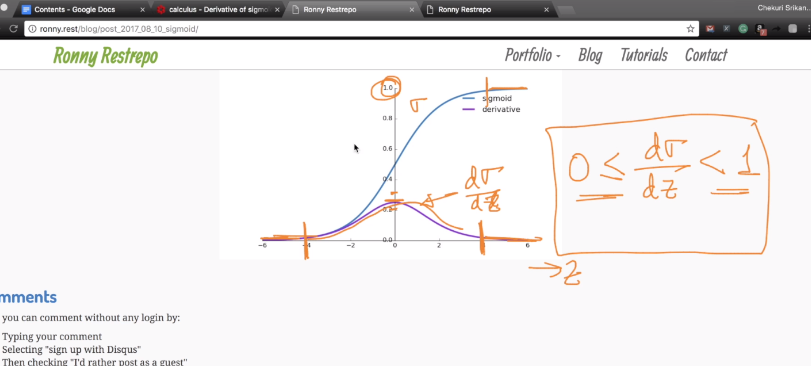
Sigmoid function is shown in below.



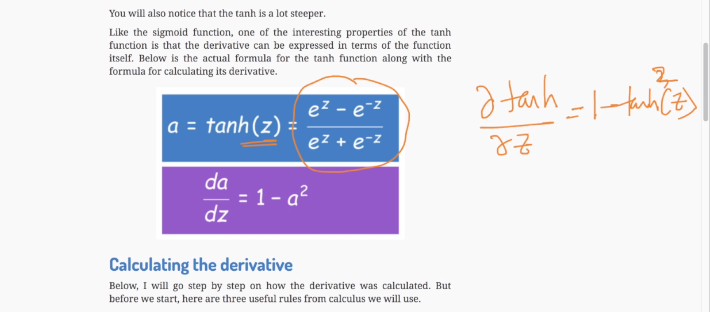
As shown in below image derivative of sigmoid can be expressed in terms of sigmoid therefore if we implement sigmoid function we can use it in forward propagation as output of function propagates forward and also we can use sigmoid function for computing derivatives.

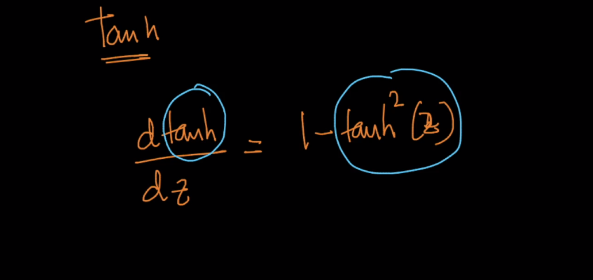


In below image it is shown sigmoid in sky blue and derivative of sigmoid in purple therefore by this it is seen that derivative of sigmoid is between 0 and 1 .



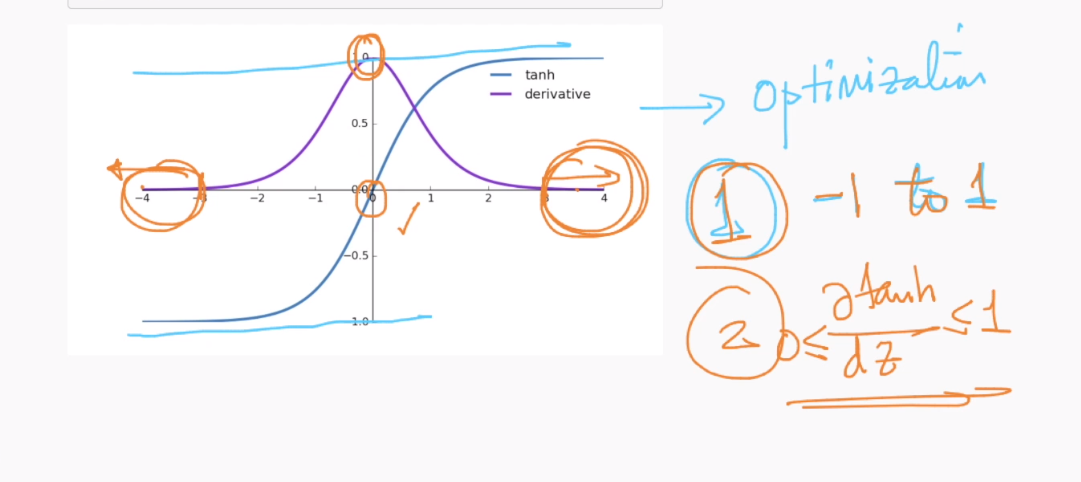
Another important activation function is tanh function. As shown below it is almost same as sigmoid function and also derivative of tanh function can be expressed in terms of tanh.





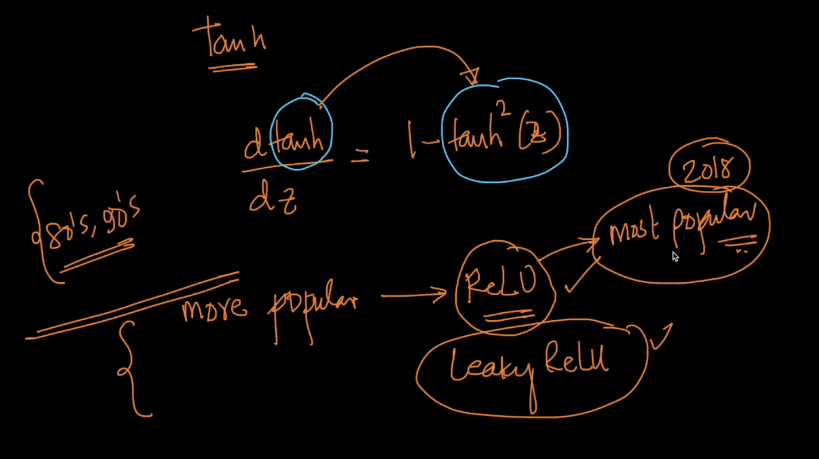
In below image light blue curve shows tanh and purple shows derivative of tanh

Curve of tanh is between -1 and 1 unlike sigmoid which is b/w 0 and 1 and also derivative of tanh is b/w o and 1





Above activation functions are old, today relu and leaky relu are most popular one



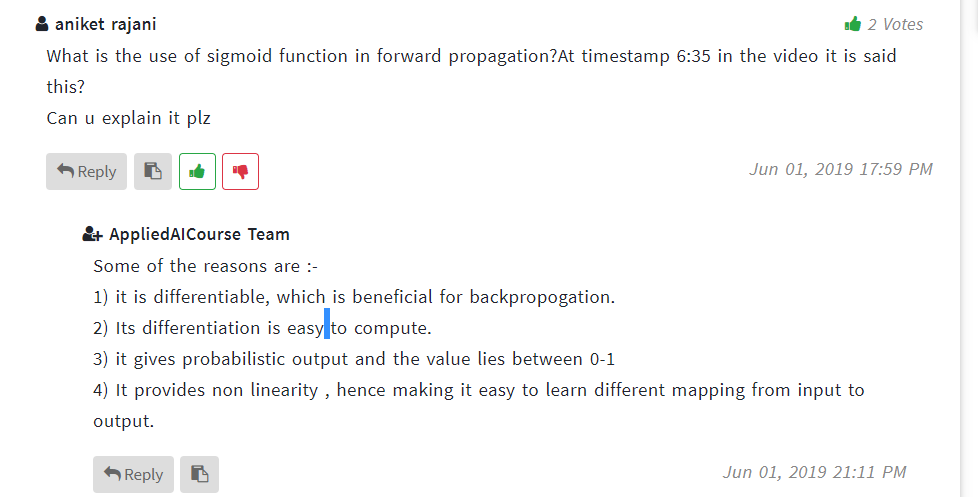
Link :

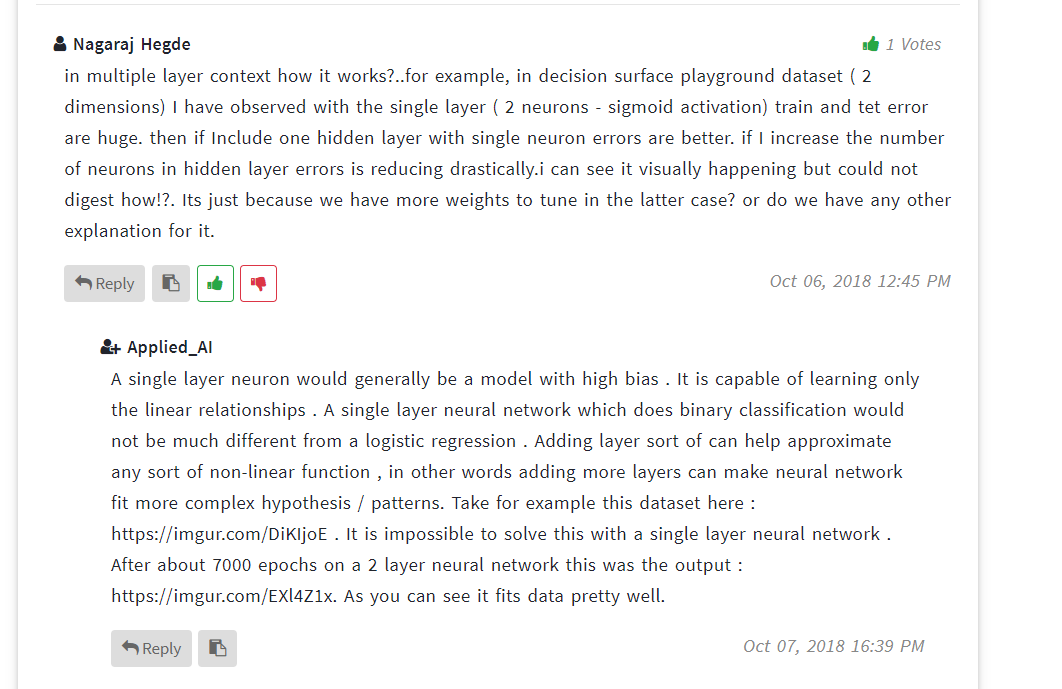
Sigmoid derivative : <https://drive.google.com/file/d/1o0B3sPwiC8aCp4a_locFfSD9753CA-tm/view>

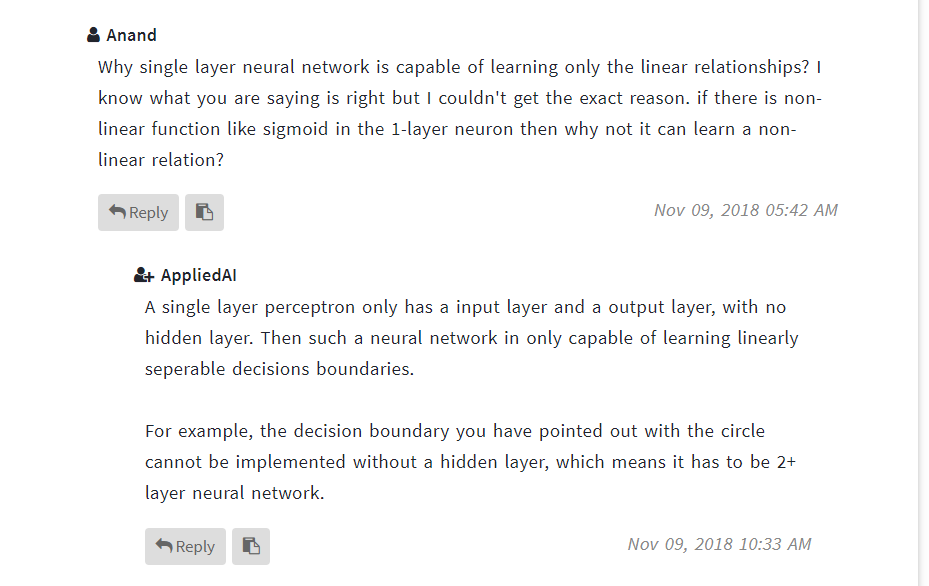
Tanh derivative : <https://drive.google.com/file/d/1aHH4uDIb5SPe3DXLd_capSREkoj77Cdn/view>

Neural network computes any function : <http://neuralnetworksanddeeplearning.com/chap4.html>

Comments :







<https://www.quora.com/Whats-the-reason-motivation-behind-using-a-sigmoid-logistic-activation-function-multiple-times-instead-of-different-activation-functions-one-at-a-time-in-a-neural-network-of-machine-learning>

