

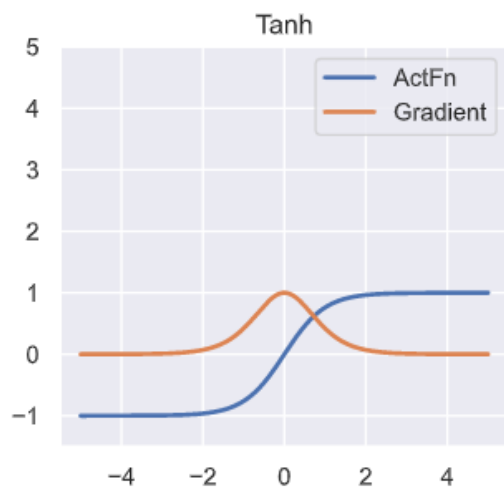
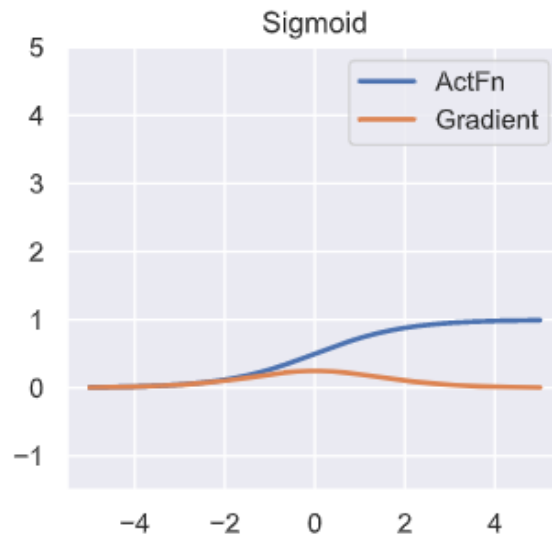


Activation Functions

The Activation Functions notebook allowed me to distinguish between six different activation functions such as Sigmoid, Tanh, ReLU, LeakyReLU, ELU and Swish. Also, it helped me to see which one performs better.

Sigmoid

This activation function doesn't work very well with Deep Neural Networks (DNN) because the gradient vanishes.

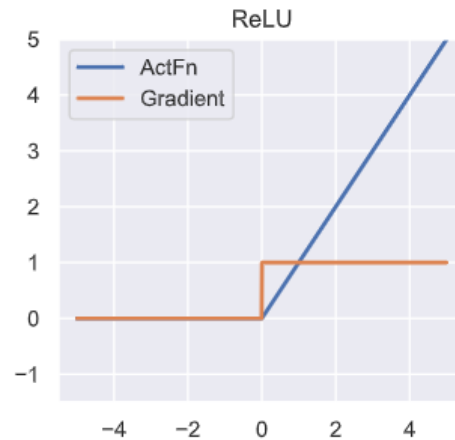


Tanh

As well as Sigmoid, this activation function doesn't work very well with DNN, but the problem is not that bad.

ReLu

This activation function works very well with DNN but sometimes the gradients go to zero due to its definition.



LeakyReLU, ELU and Swish

These three activation functions help solving ReLU's problem because their definitions of negative x is not zero.

