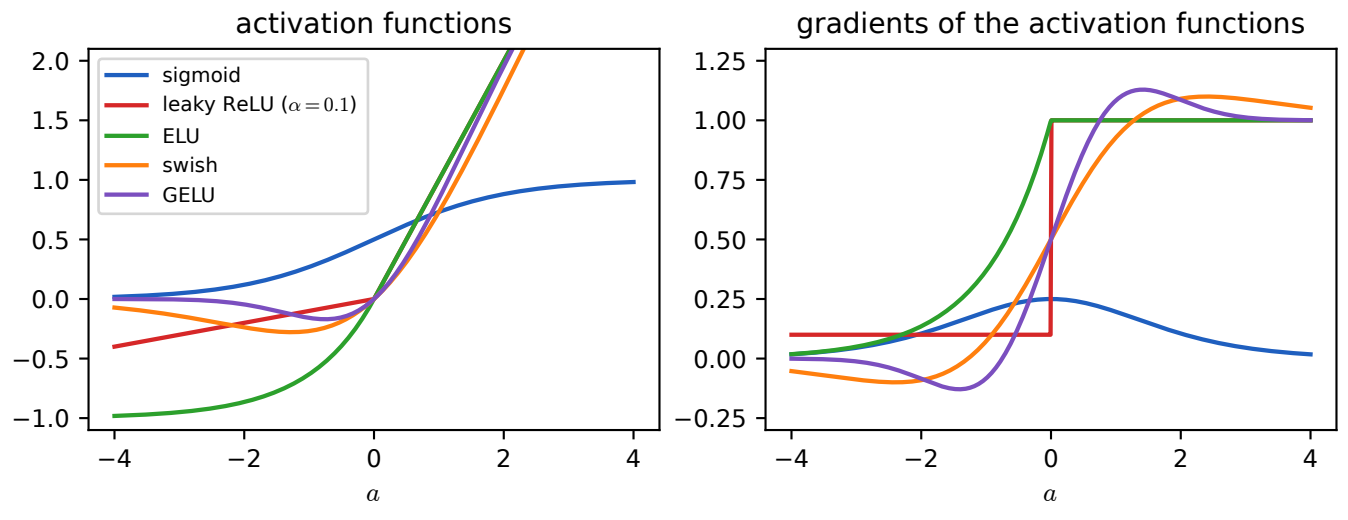

activation functions
gradients of the activation functions
sigmoid
leaky ReLU ($\alpha = 0.1$)
ELU
swish
GELU
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1.25
1.00
0.75
0.50
0.25
0.00
−0.25
−4
−2
0
2
4
a