

$$\text{MLP: } f_{\theta}(x) = A_L \circ \sigma \circ A_{L-1} \dots \circ \sigma \circ A_1(x), \theta = (A_1, \dots, A_L)$$