Multilayer neural network output z output L hidden layers w_{ki} Layer pre-activation for i>0, lo=X, h(o)=L $h_{(i)} = W_{(i-1)} * L_{(i-1)}$ hidden Hidden layer activation (from 1 1 to L) W_{jj} $h_{(1)} = a(h_{(1)})$ input input x