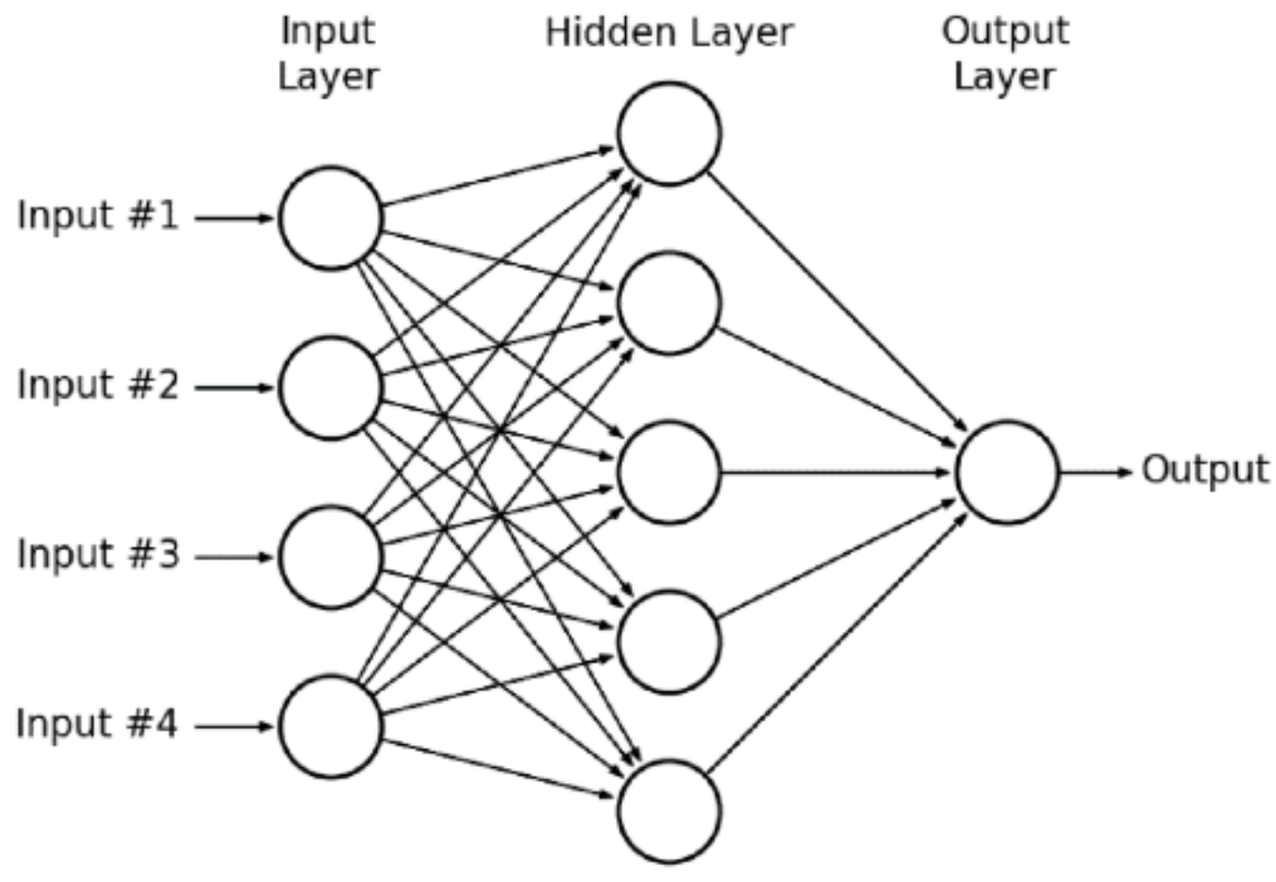
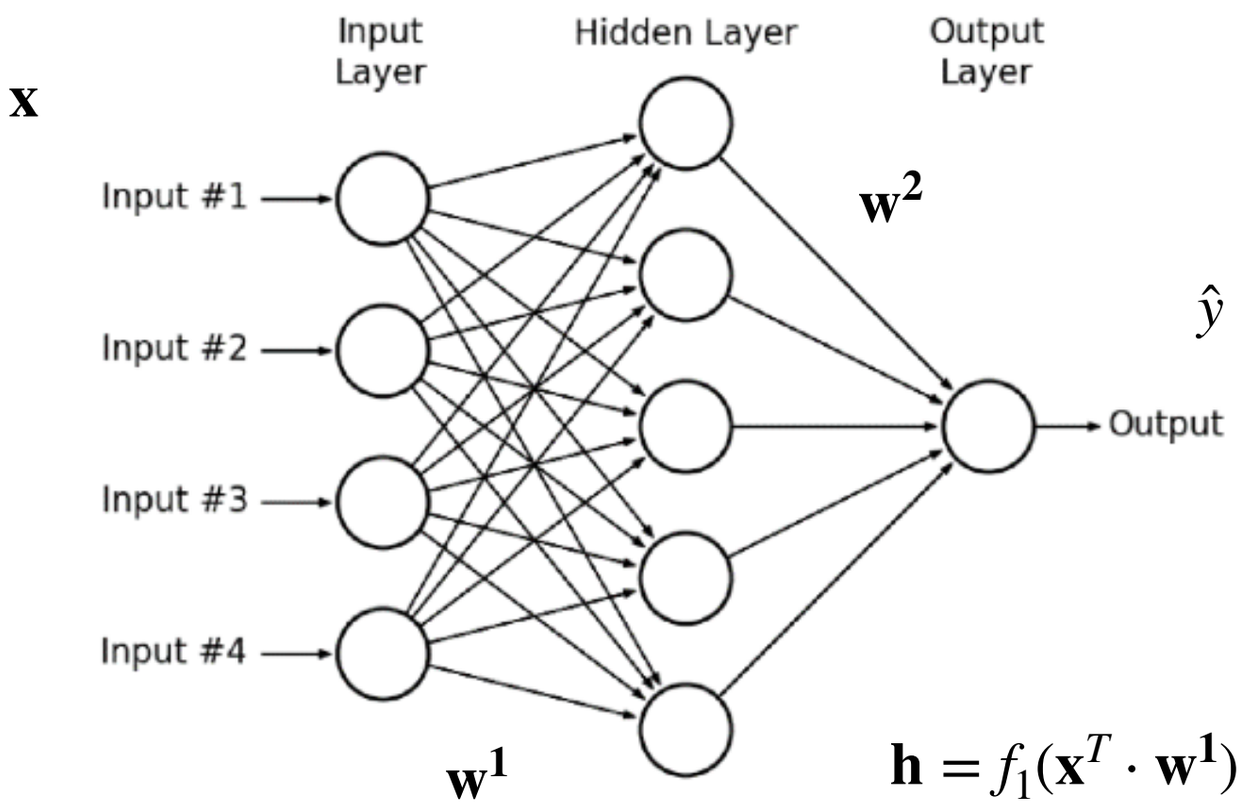
Training an MLP

Backpropagation



Backpropagation

$$l = \frac{1}{2}(\hat{y} - y)^2$$



Assuming MSE loss function and linear activation function at the output

$$\frac{\partial l}{\partial w^2} = \frac{\partial l}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial w^2} = (\hat{y} - y)h$$

$$\frac{\partial l}{\partial w^1} = \frac{\partial l}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial h} \frac{\partial h}{\partial w^1} = (\hat{y} - y)w^2 f_1'(w^1 x)x$$