Training an MLP

Gradient descent on the perceptron

- Compute the output, $\hat{y}$

- Compute the gradient of the loss function wrt the parameters, $\dfrac{\partial l}{\partial w}$

- Update parameters, $w \leftarrow w - \eta \dfrac{\partial l}{\partial w}$

- Repeat until convergence

Training an MLP

Backpropagation