Back Propagation

Heni Ben Amor, Ph.D.
Assistant Professor
Arizona State University

Ira A. Fulton Schools of Engineering
Arizona State University

Approach

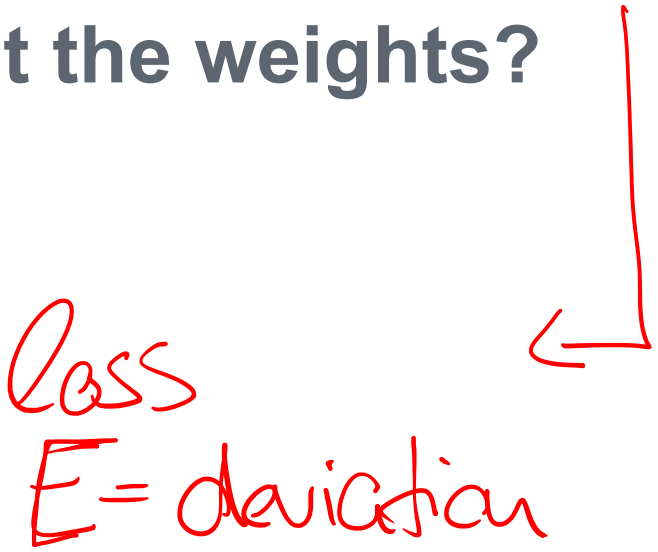
- Given a set of training data
- Each sample a tuple $< \mathbf{x}, \mathbf{y} >$
- Where $\mathbf{x}$ is the input and $\mathbf{y}$ is the desired output
- Train network such that

$$NN(\mathbf{x}) \approx \mathbf{y} \quad \forall \mathbf{x} \in \mathbf{X}$$

- Assumes labeled training data
- Typically labels are provided by human annotation

Back Propagation Algorithm

- Algorithm for training the weights
- Start with random weights
- Adjust the weights to reduce error
- How do we adjust the weights?
- Random search?



Back Propagation (BP)

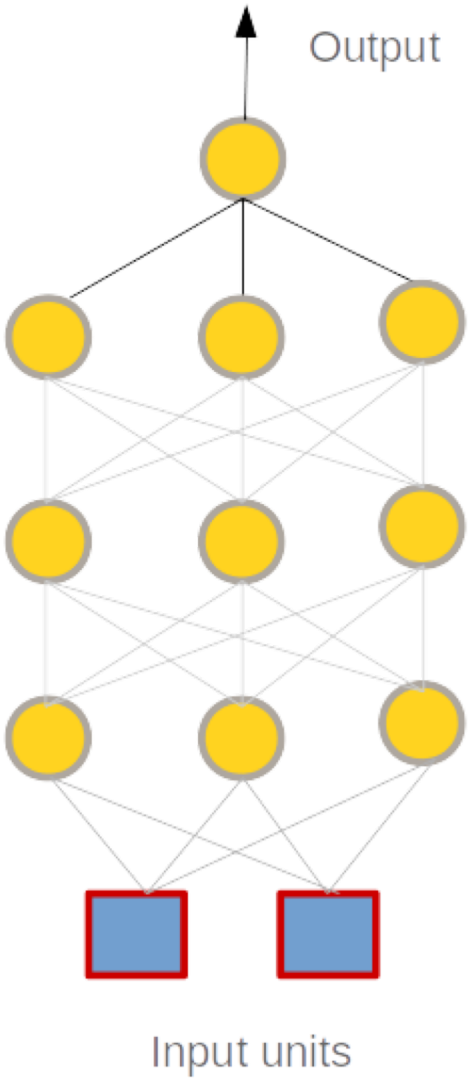
- Given input and output, learn **weights**
- Gradient descent minimizing quadratic error:

$$E = \frac{1}{2} \sum_{i=1}^{N} ||y_i - \hat{y}_i||^2$$

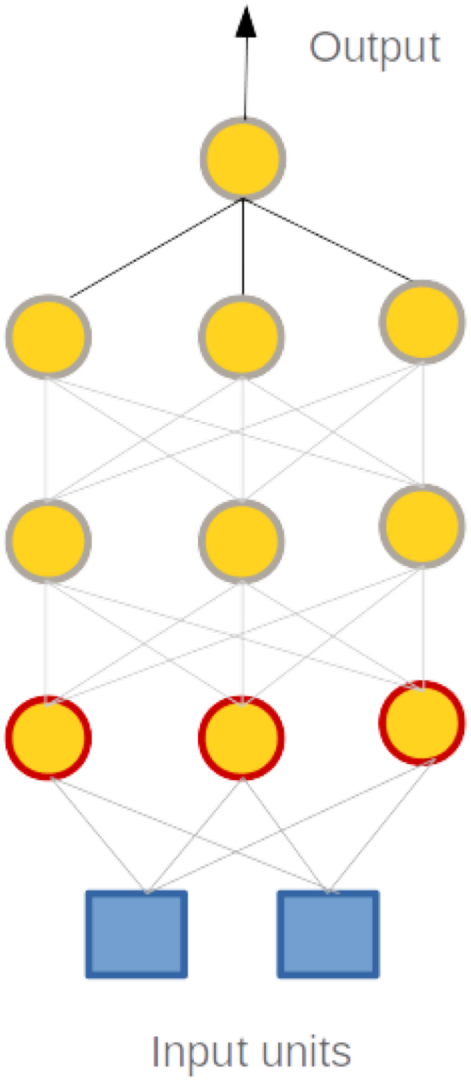
target — output of ANN

- In other words, minimize quadratic difference between target and output of the network

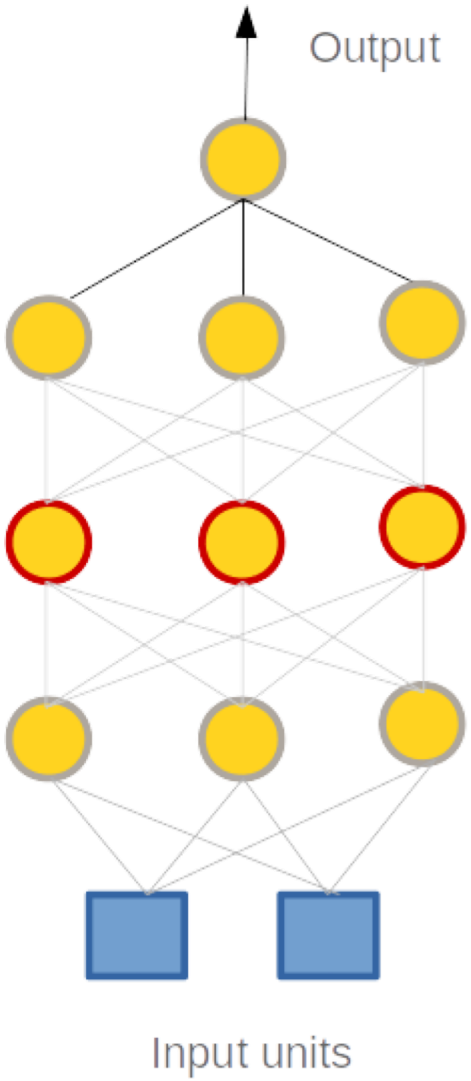
Forward Propagation



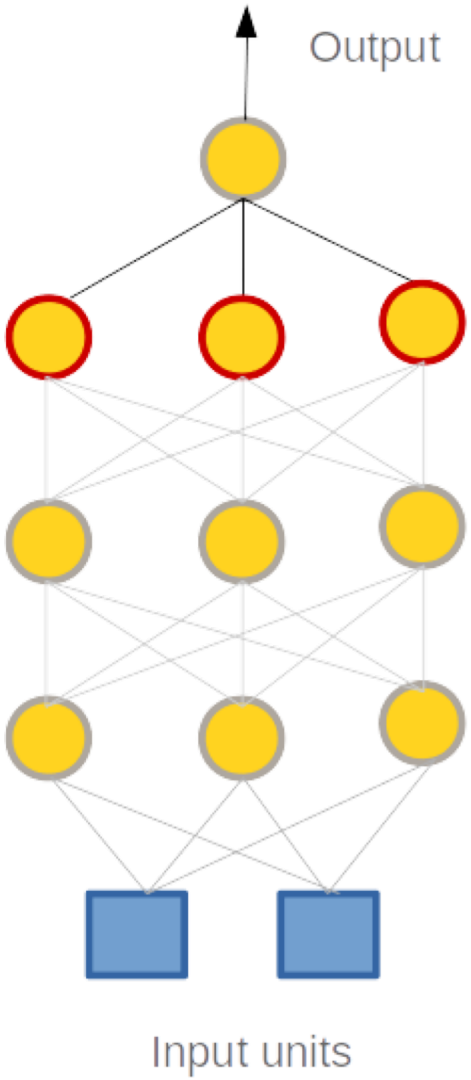
Forward Propagation



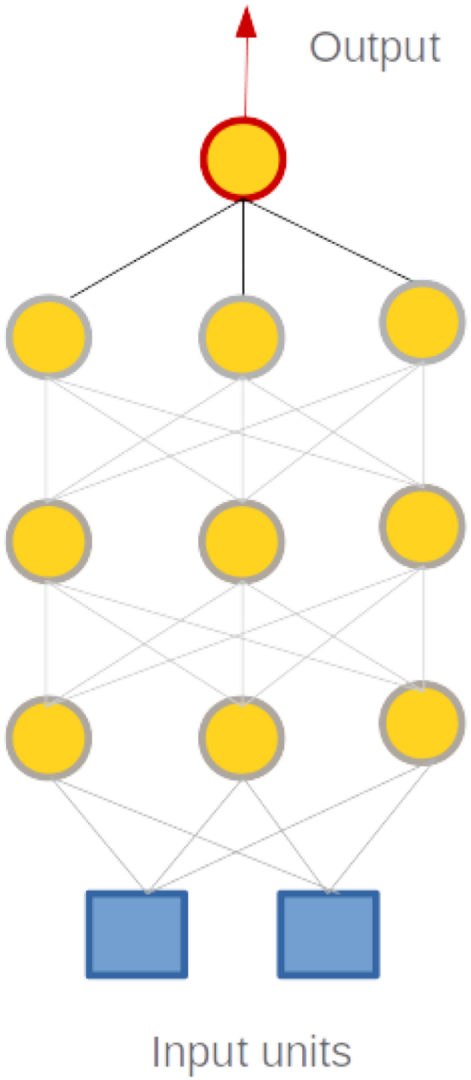
Forward Propagation



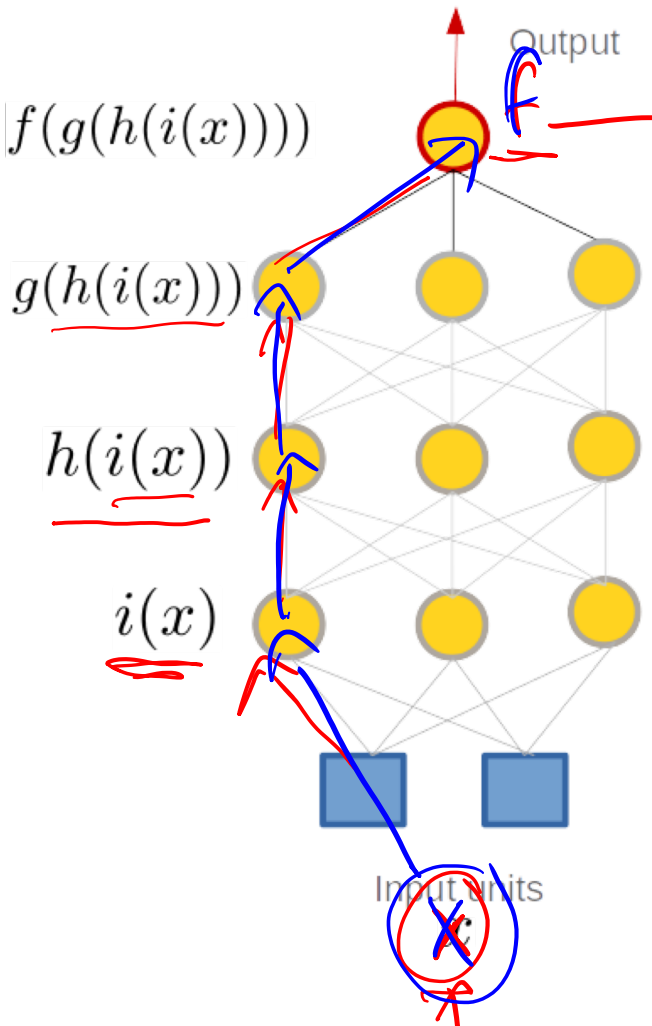
Forward Propagation



Forward Propagation



Neural Networks as Nested Functions



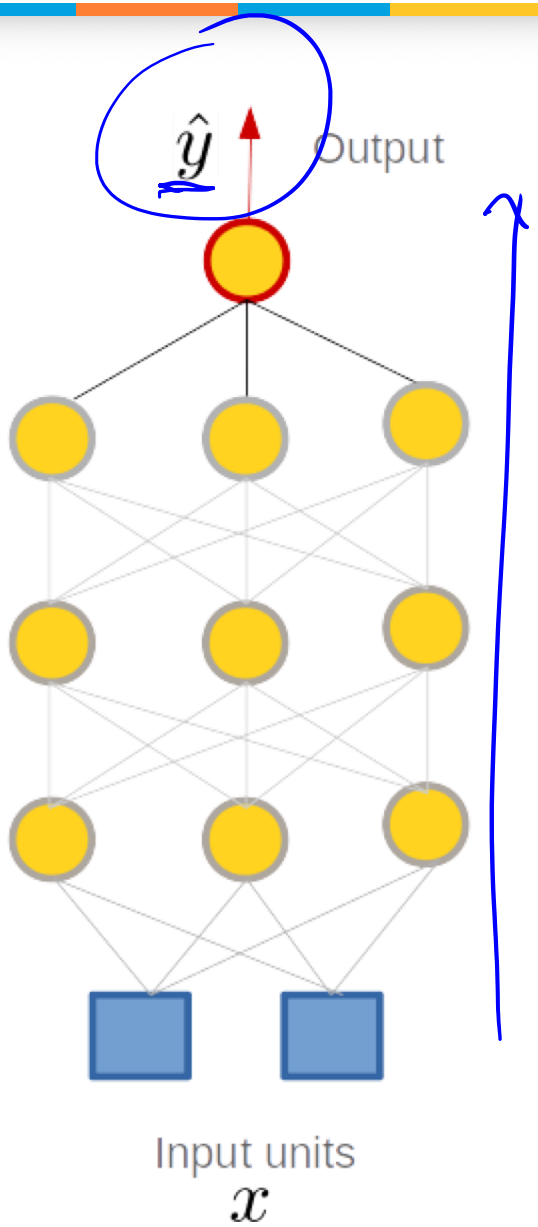
Neural Network as a Nested Function

$$f(g(h(i(x))))$$

Calculating the partial derivatives via chain rule

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial g}\frac{\partial g}{\partial h}\frac{\partial h}{\partial i}\frac{\partial i}{\partial x}$$

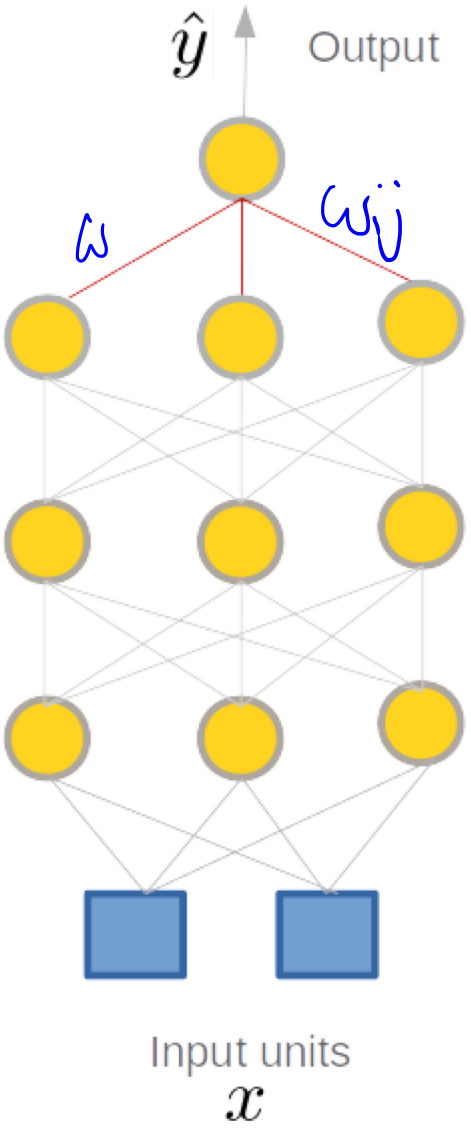
Intuition of Back Propagation



1. **Calculate Error at Final Layer**

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

Intuition of Back Propagation

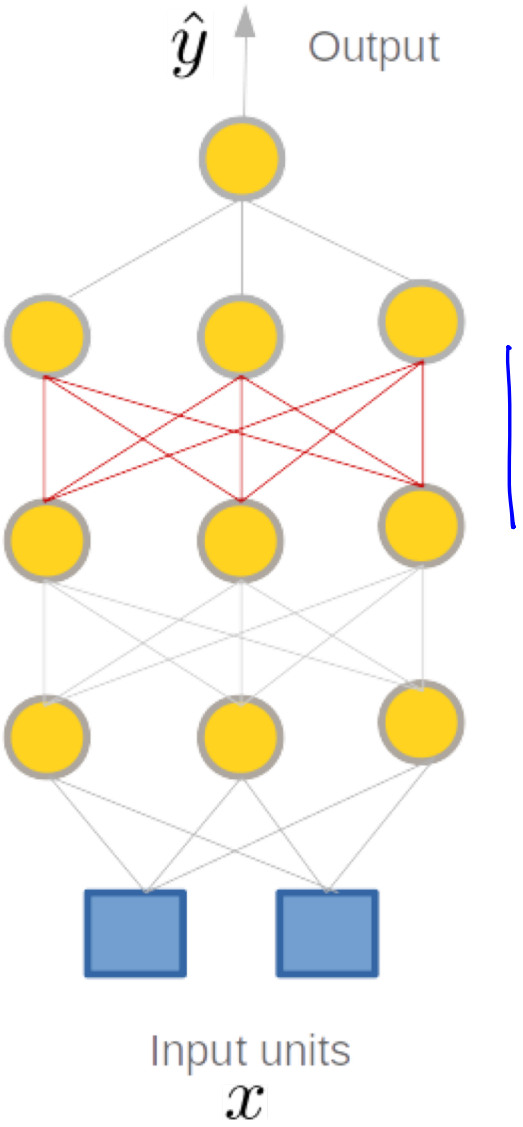


1. **Calculate Error at Final Layer**

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

2. **Go back through layers and calculate changes to all weights**

Intuition of Back Propagation



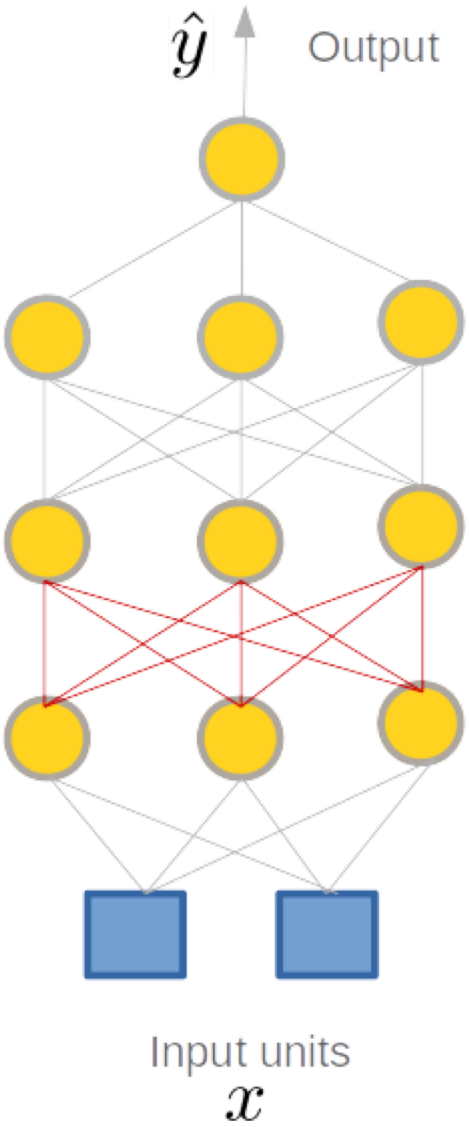
1. Calculate Error at Final Layer

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

2. Go back through layers and calculate changes to all weights

3. Repeat step 2 for all previous layers

Intuition of Back Propagation



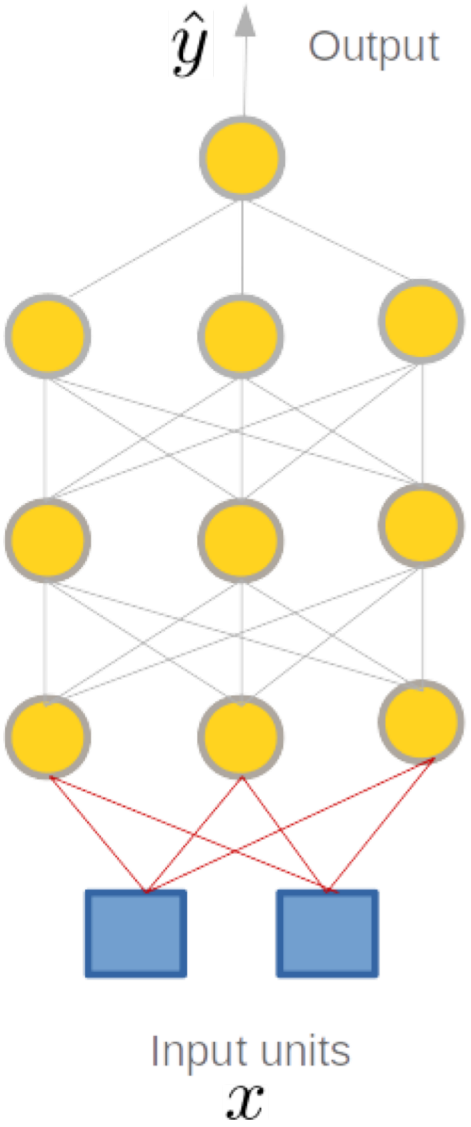
1. Calculate Error at Final Layer

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

2. Go back through layers and calculate changes to all weights

3. Repeat step 2 for all previous layers

Intuition of Back Propagation



1. Calculate Error at Final Layer

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

2. Go back through layers and calculate changes to all weights

3. Repeat step 2 for all previous layers

Summary

- Back propagation algorithm
- Generates the weights of the network
- Objective is to minimize the training error
- Perform gradient descent
- Calculate derivative of network
- Derivatives are propagated backwards