Back-Propagation Neural Networks

Overview

- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- Hidden Layer Networks
- Training with BP
- Examples

- Outline to the left
- Current topic in red
  - Introduction
  - History and Origins
  - Biologically Inspired
  - Applications
  - Perceptron
  - Activation Functions
  - Hidden Layer Networks
  - Training with BP
  - Examples

Introduction

- Artificial Neural Networks (ANN)
  - Connectionist computation
  - Parallel distributed processing
  - Computational models
- Biologically Inspired computational models
- Machine Learning
- Artificial intelligence

History

- McCulloch and Pitts introduced the Perceptron in 1943.
  - Simplified model of a biological neuron
- Fell out of favor in the late 1960's
  - (Minsky and Papert)
  - Perceptron limitations
- Resurgence in the mid 1980's
  - Nonlinear Neuron Functions
  - Back-propagation training

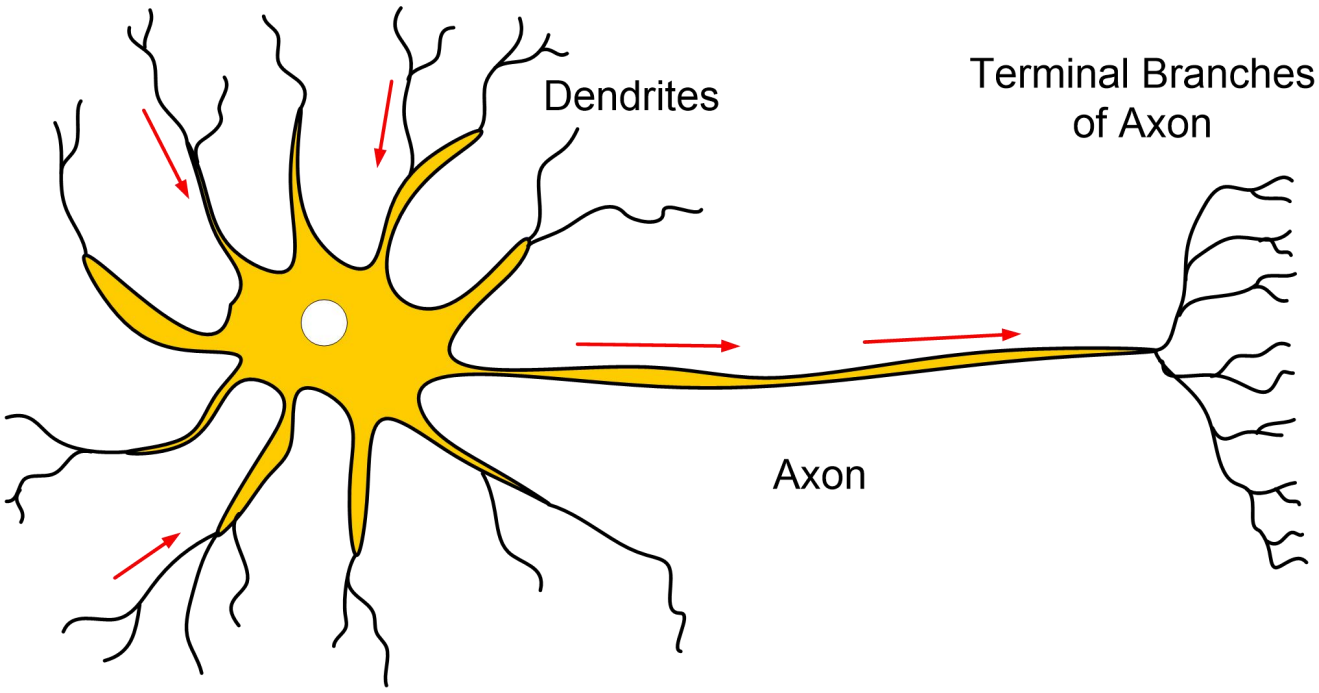
Summary of Applications

- Function approximation
- Pattern recognition
- Signal processing
- Modeling
- Control
- Machine learning

Biologically Inspired

- Introduction
- History
- **Biologically Inspired**
- Applications
- The Perceptron
- Activation Functions
- Hidden Layer Networks
- Training with BP
- Examples

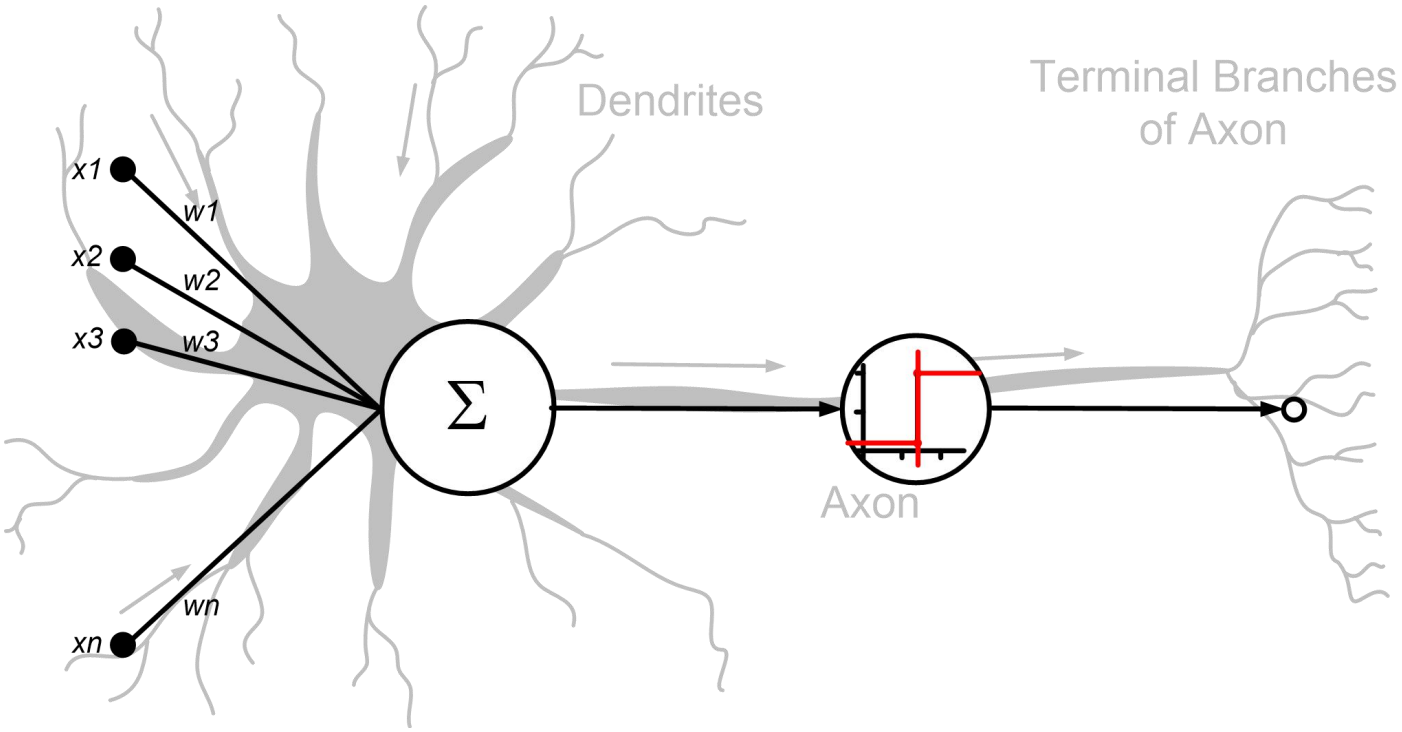
- Electro-chemical signals
- Threshold output firing



The Perceptron

- Introduction
- History
- Biologically Inspired
- Applications
- **The Perceptron**
- Activation Functions
- Hidden Layer Networks
- Training with BP
- Examples

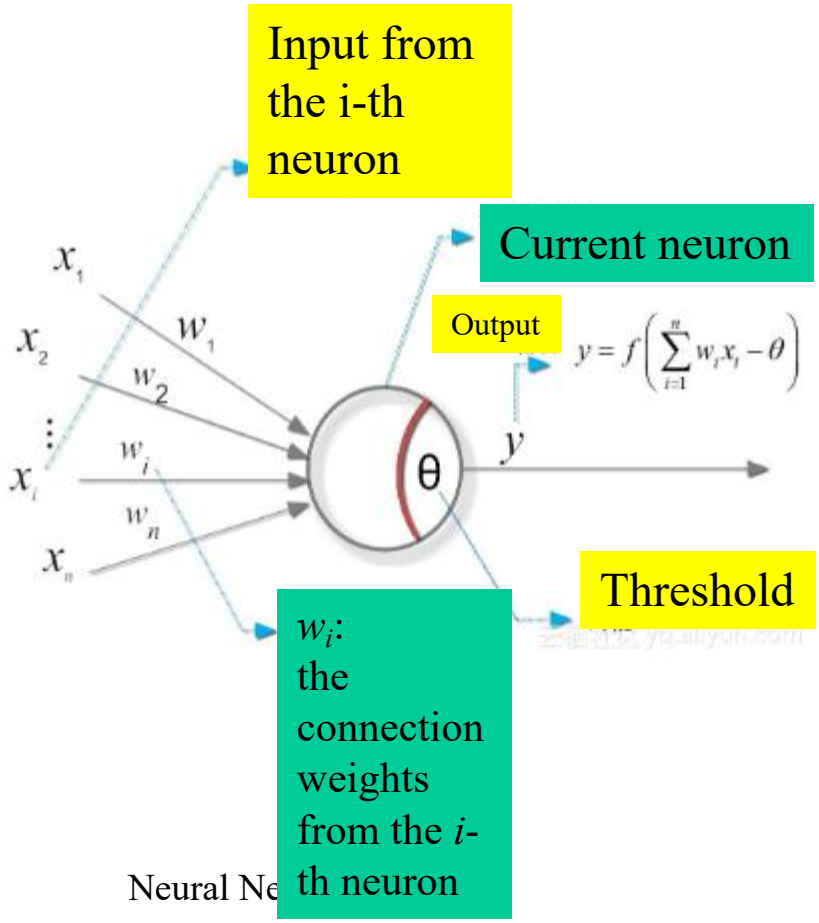
- Binary classifier functions
- Threshold activation function



The Perceptron

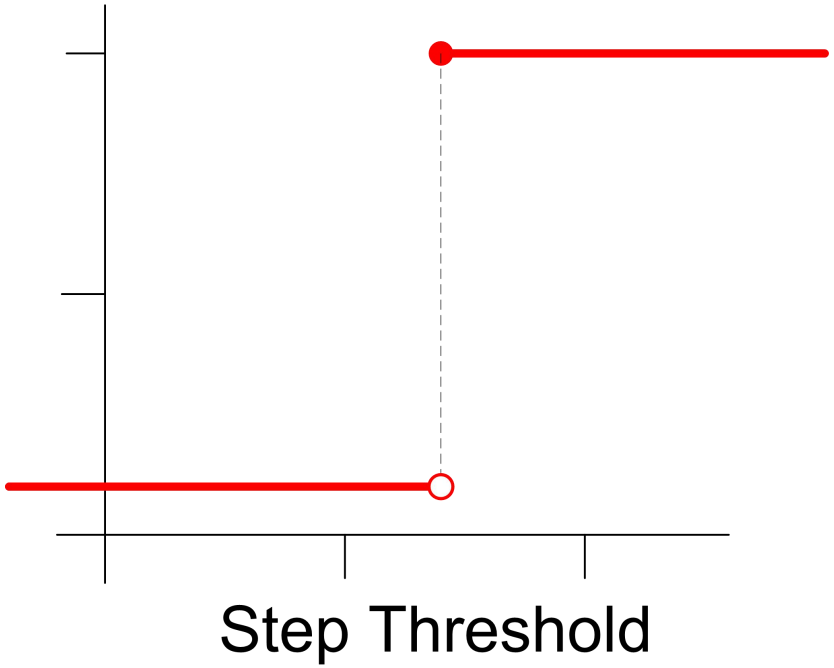
- Introduction
- History
- Biologically Inspired
- Applications
- **The Perceptron**
- Activation Functions
- Hidden Layer Networks
- Training with BP
- Examples

- Binary classifier functions
- Threshold activation function



The Perceptron: Threshold Activation Function

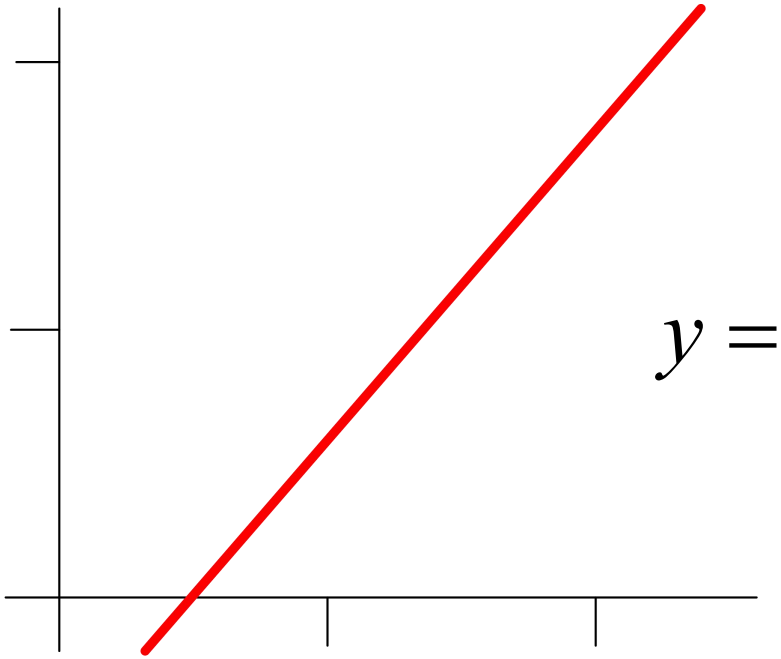
- Binary classifier functions
- Threshold activation function



Linear Activation functions

- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- **Activation Functions**
- Hidden Layer Networks
- Training with BP
- Examples

- Output is scaled sum of inputs



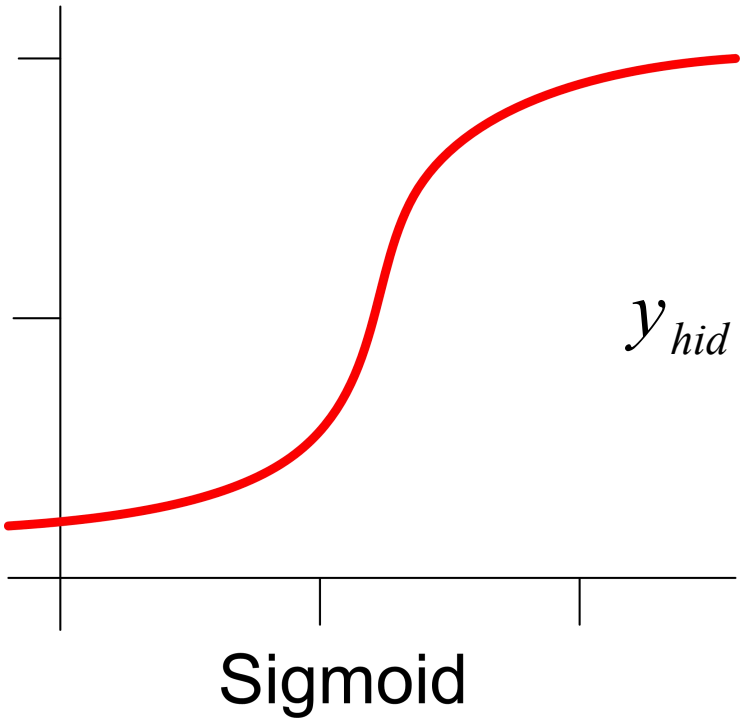
Linear

$$y = u = \sum_{n=1}^{N} w_n x_n$$

Nonlinear Activation Functions

- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- **Activation Functions**
- Hidden Layer Networks
- Training with BP
- Examples

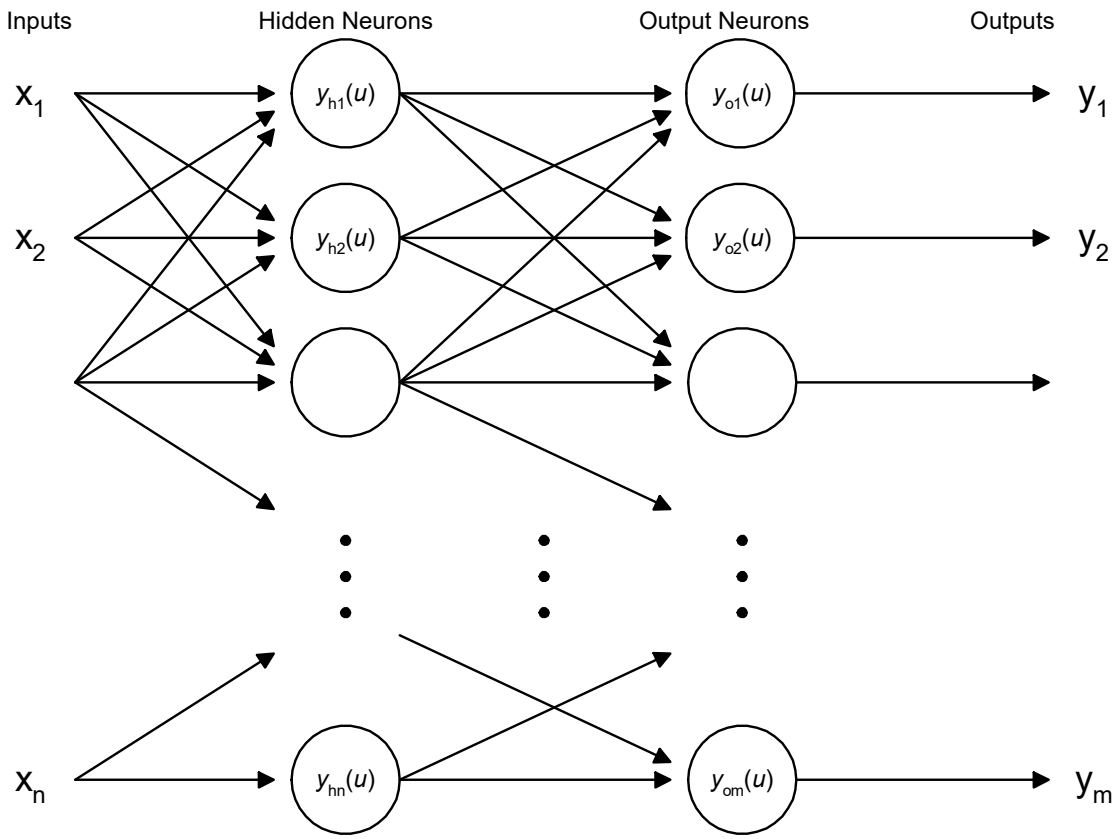
- Sigmoid Neuron unit function



$$y_{hid}(u) = \frac{1}{1+e^{-u}}$$

Layered Networks

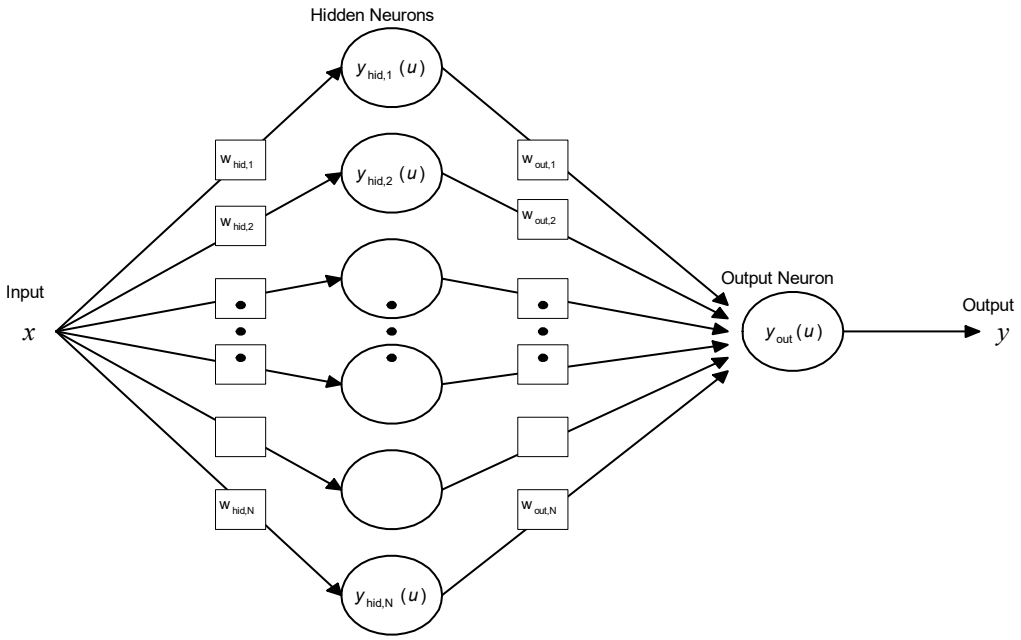
- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- **Hidden Layer Networks**
- Training with BP
- Examples



SISO Single Hidden Layer Network

- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- **Hidden Layer Networks**
- Training with BP
- Examples

- Can represent and single input single output functions: $y = f(x)$

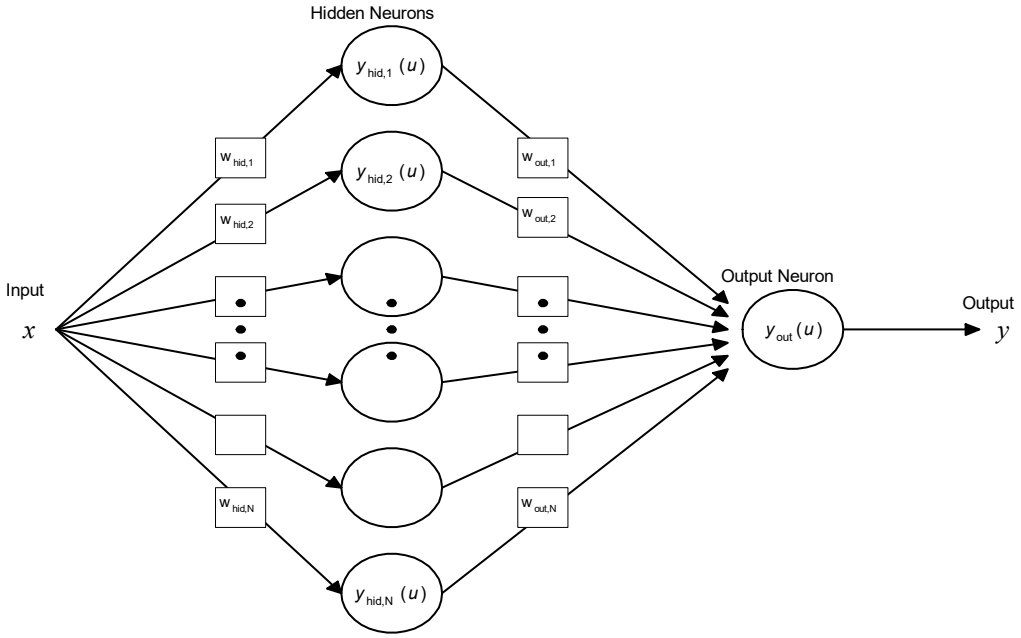


Training Data Set

- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- **Hidden Layer Networks**
- Training with BP
- Examples

Adjust weights (w) to learn a given target function: $y = f(x)$

Given a set of training data $X \rightarrow Y$



Training Weights: Error Back-Propagation (BP)

- Weight update formula:

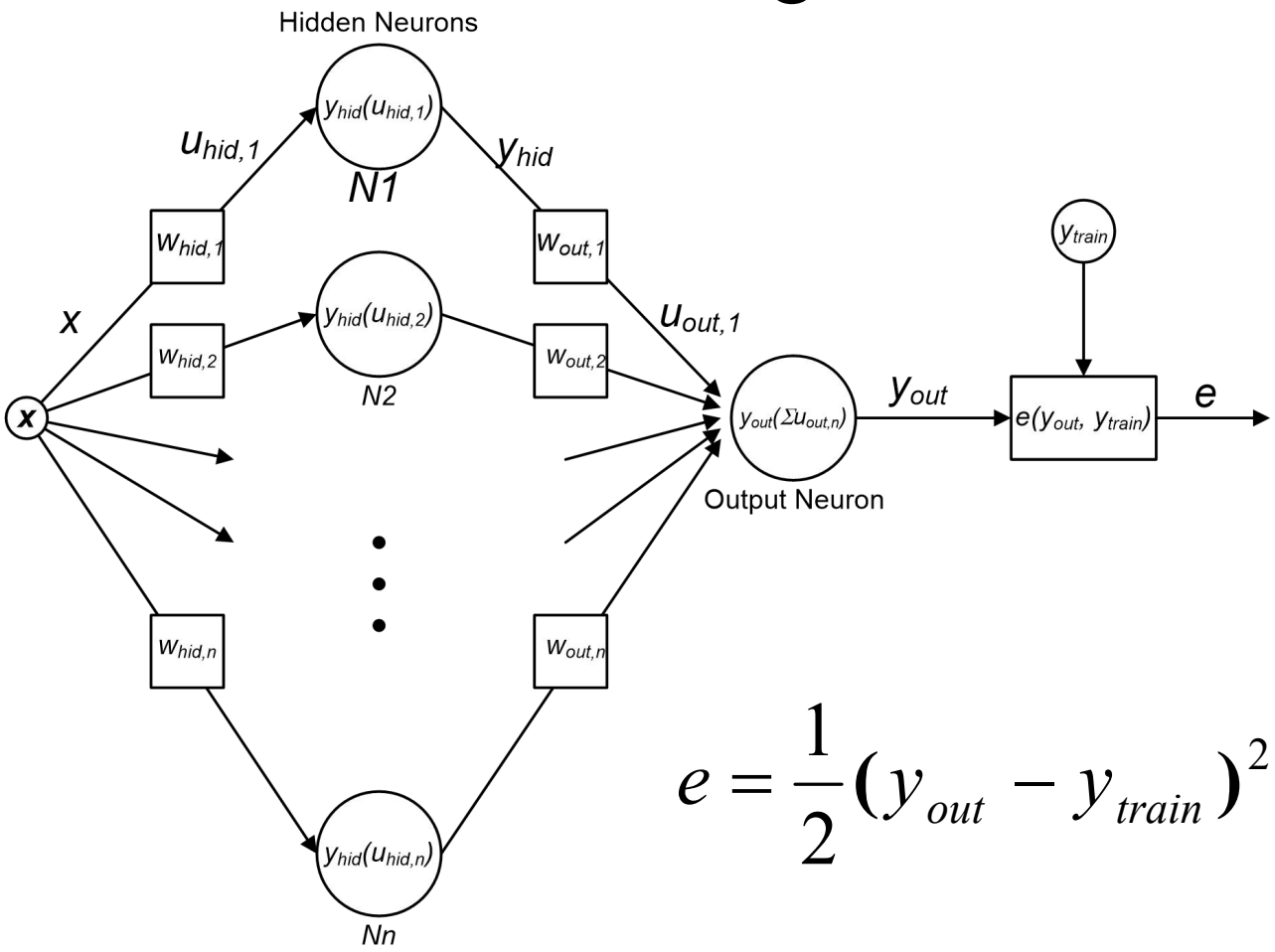
$$w(k+1) = w(k) + \Delta w$$

$$\Delta w(i) = \eta * \frac{\partial e(i)}{\partial w}$$

Error Back-Propagation (BP)

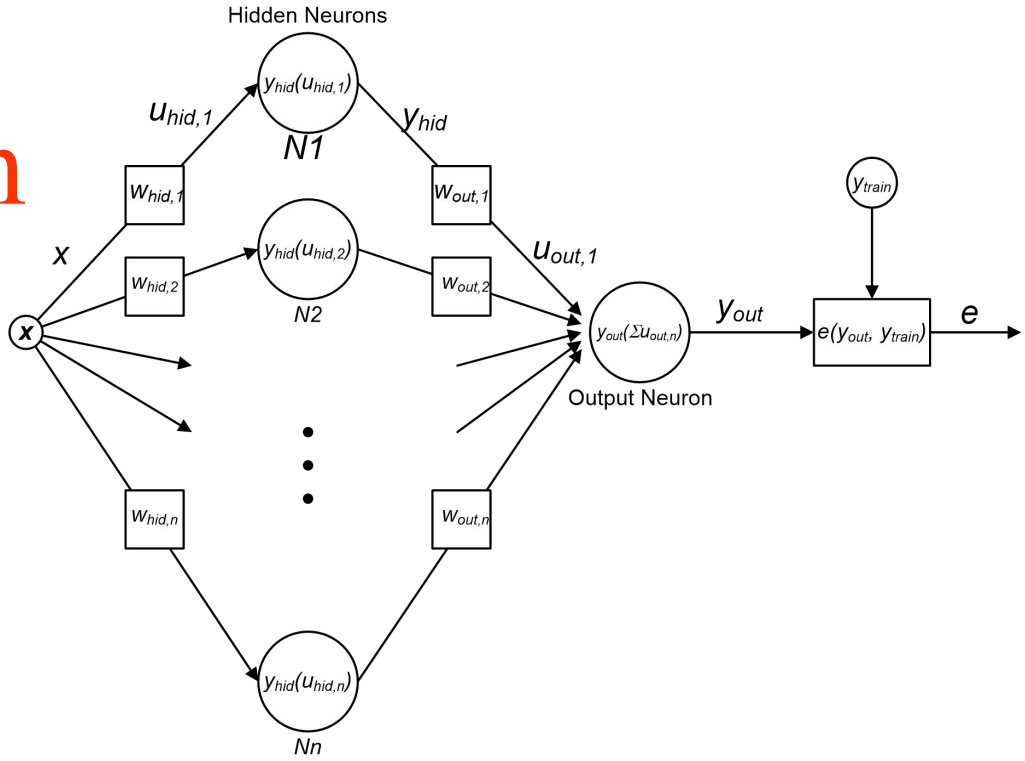
- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- Hidden Layer Networks
- **Training with BP**
- Examples

Training error term: e



$$e = \frac{1}{2}(y_{out} - y_{train})^2$$

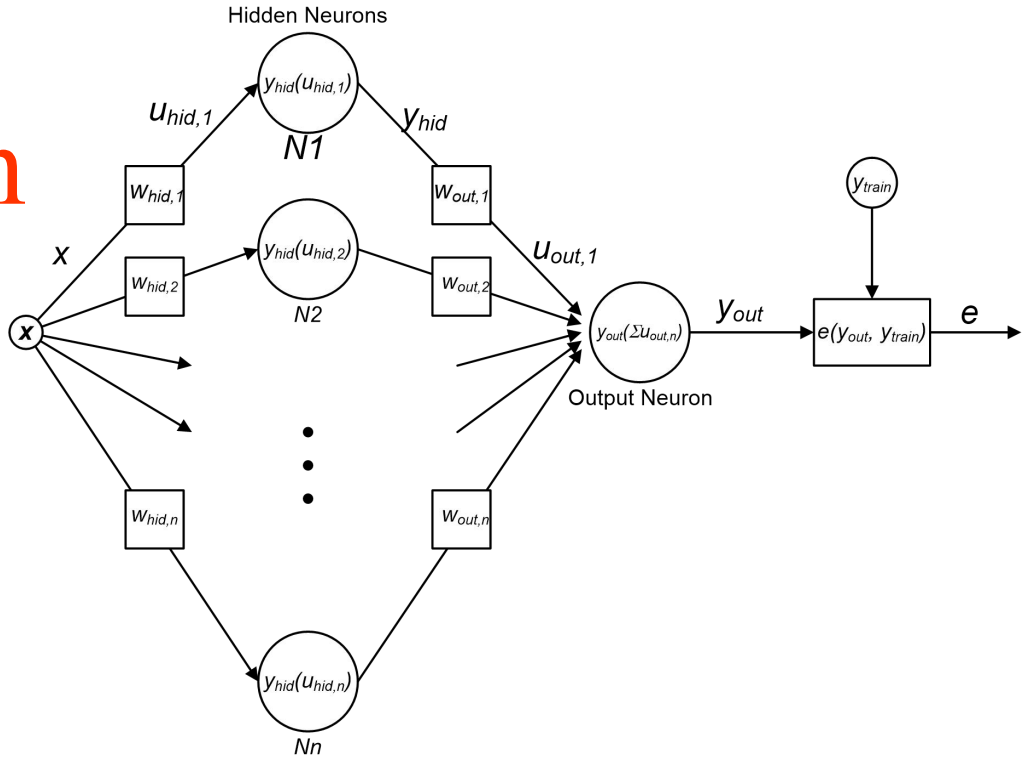
BP Formulation



- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- Hidden Layer Networks
- **Training with BP**
- Examples

$$e(y_{out}, y_{train}) = e(y_{out}(u_{out,1}), y_{train})$$
$$= e(y_{out}(w_{out,1} y_{hid,1}), y_{train})$$
$$= e(y_{out}(w_{out,1} y_{hid}(u_{hid,1})), y_{train})$$
$$= e(y_{out}(w_{out,1} y_{hid}(w_{hid,1} x)), y_{train})$$

BP Formulation



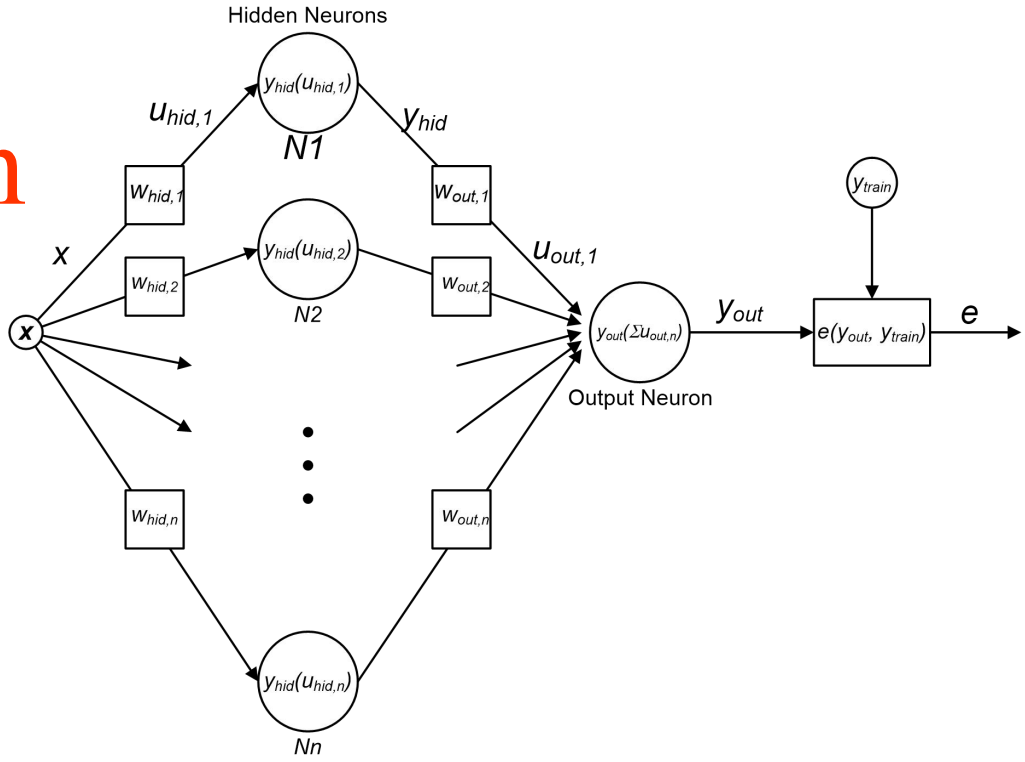
- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- Hidden Layer Networks
- **Training with BP**
- Examples

$$\frac{\partial e}{\partial w_{hid}} = \frac{\partial e}{\partial y_{out}} \frac{\partial y_{out}}{\partial u_{out,1}} \frac{\partial u_{out,1}}{\partial y_{hid}} \frac{\partial y_{hid}}{\partial u_{hid,1}} \frac{\partial u_{hid,1}}{\partial w_{hid,1}}$$

$$\frac{\partial e}{\partial w_{hid}} = \frac{\partial u_{hid,1}}{\partial w_{hid,1}} \frac{\partial y_{hid}}{\partial u_{hid,1}} \frac{\partial u_{out,1}}{\partial y_{hid}} \frac{\partial y_{out}}{\partial u_{out,1}} \frac{\partial e}{\partial y_{out}}$$

BP Formulation

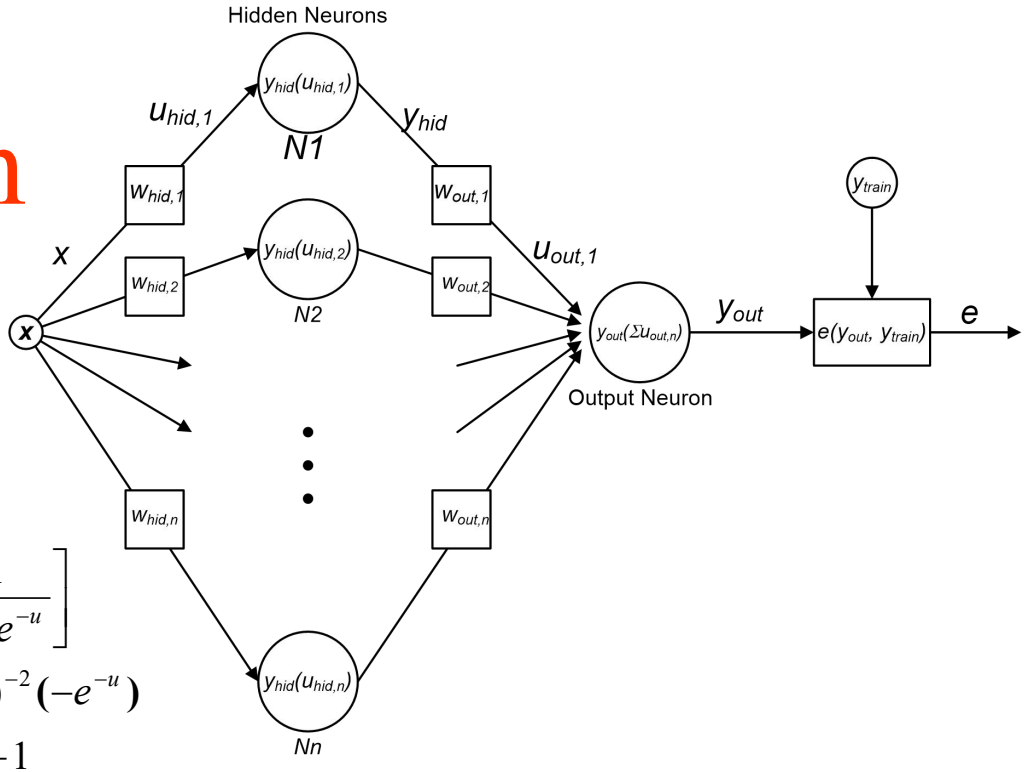
- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- Hidden Layer Networks
- **Training with BP**
- Examples



$$\frac{\partial u_{hid,1}}{\partial w_{hid,1}} = \frac{\partial}{\partial w_{hid,1}} w_{hid,1} x$$

$$= x$$

BP Formulation



- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- Hidden Layer Networks
- **Training with BP**
- Examples

$$\frac{dy_{hid}(u)}{du} = \frac{d}{du}\left[\frac{1}{1+e^{-u}}\right]$$

$$= (1+e^{-u})^{-2}(-e^{-u})$$

$$= \frac{1+e^{-u}-1}{(1+e^{-u})^2}$$

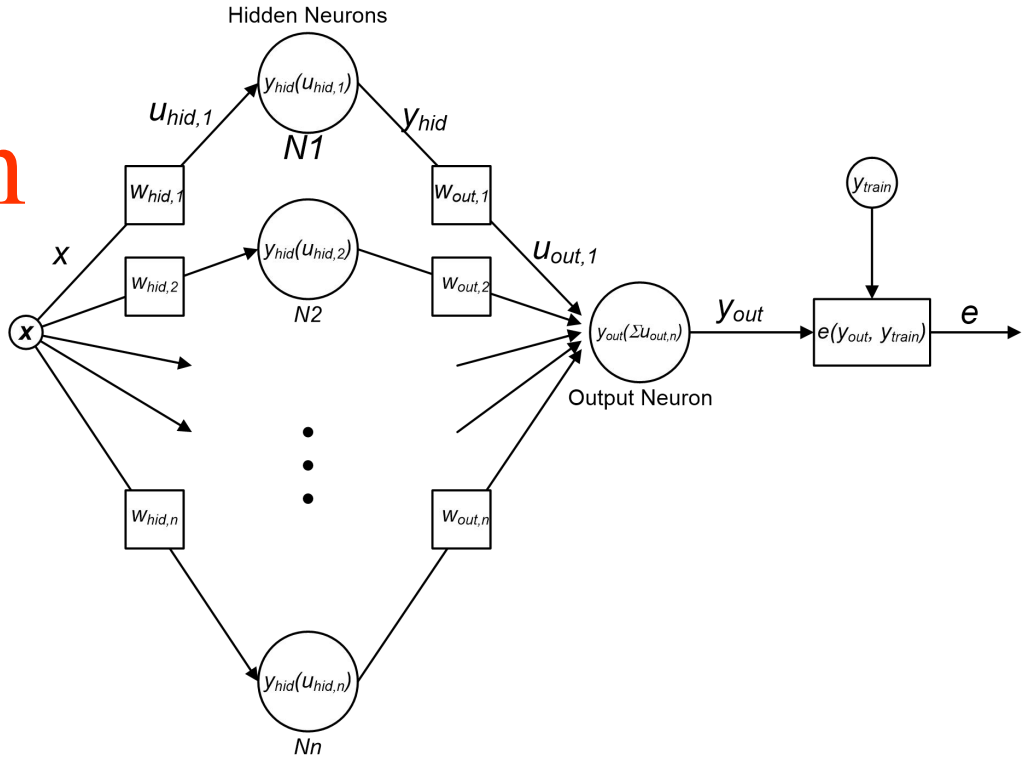
$$= \frac{1+e^{-u}}{(1+e^{-u})^2} - \frac{1}{(1+e^{-u})^2}$$

$$= \frac{1}{(1+e^{-u})} - \frac{1}{(1+e^{-u})^2}$$

$$= \frac{1}{(1+e^{-u})}\left[1 - \frac{1}{(1+e^{-u})}\right]$$

$$= y_{hid}(u_{hid,1})\left[1 - y_{hid}(u_{hid,1})\right]$$

BP Formulation



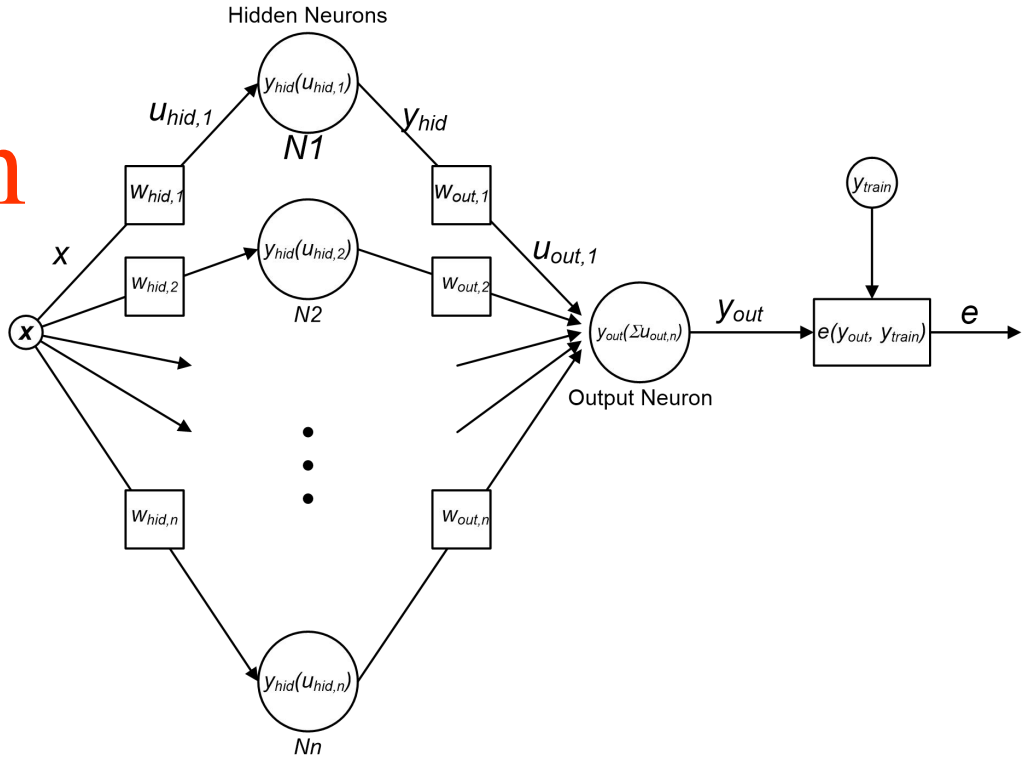
- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- Hidden Layer Networks
- **Training with BP**
- Examples

$$\frac{\partial u_{out,1}}{\partial y_{hid}} = \frac{\partial}{\partial y_{hid}} w_{out,1} y_{hid,1}$$

$$= w_{out,1}$$

BP Formulation

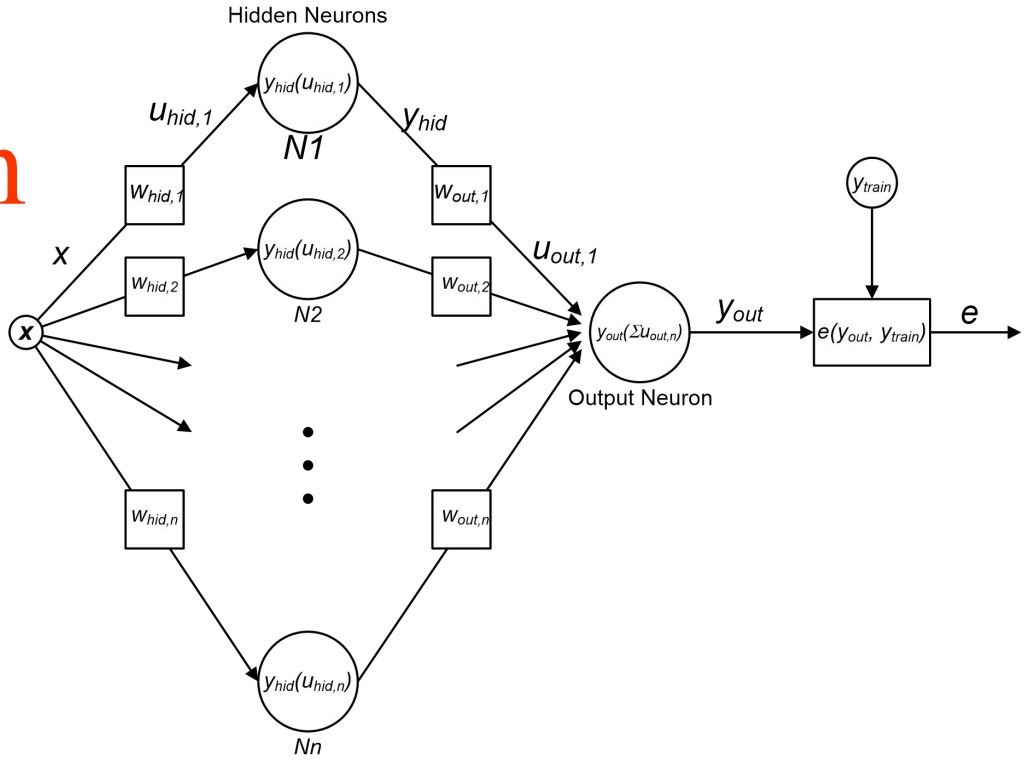
- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- Hidden Layer Networks
- **Training with BP**
- Examples



$$\frac{\partial e}{\partial y_{out}} = \frac{\partial}{\partial y_{out}} \frac{1}{2}(y_{out} - y_{train})^2$$

$$= (y_{out} - y_{train})$$

BP Formulation



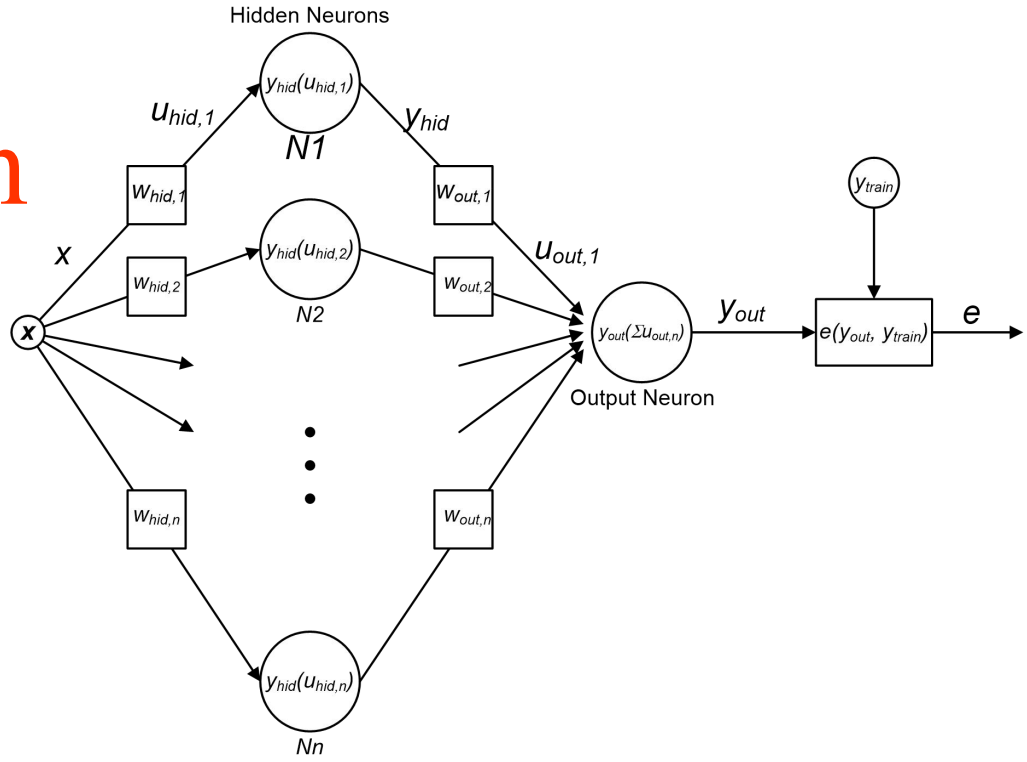
- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- Hidden Layer Networks
- **Training with BP**
- Examples

$$\frac{\partial e}{\partial w_{hid}} = \frac{\partial u_{hid,1}}{\partial w_{hid,1}} \frac{\partial y_{hid}}{\partial u_{hid,1}} \frac{\partial u_{out,1}}{\partial y_{hid}} \frac{\partial y_{out}}{\partial u_{out,1}} \frac{\partial e}{\partial y_{out}}$$

$$= (x)\left(y_{hid}(u_{hid,1})\left[1 - y_{hid}(u_{hid,1})\right]\right)(w_{out,1})(1)(y_{out} - y_{train})$$

BP Formulation

- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- Hidden Layer Networks
- **Training with BP**
- Examples



$$\frac{\partial e}{\partial w_{out}} = \frac{\partial u_{out,1}}{\partial w_{out}} \frac{\partial y_{out}}{\partial u_{out,1}} \frac{\partial e}{\partial y_{out}}$$

$$= \left(\frac{\partial}{\partial w_{out}} w_{out,1} y_{hid,1} \right) \left(\frac{\partial}{\partial u_{out,1}} \left[u_{out,1} + u_{out,2} + \ldots + u_{out,N} \right] \right) \left(\frac{\partial}{\partial y_{out}} \frac{1}{2} (y_{out} - y_{train})^2 \right)$$

$$= (y_{hid})(1)(y_{out} - y_{train})$$

Example: Step by step

- Introduction
- History
- Biologically Inspired
- Applications
- The Perceptron
- Activation Functions
- Hidden Layer Networks
- Training with BP
- **Examples**