Exercise

Animal	Mammal?	Can fly?	Carnivore?
Bat	Yes	Yes	Yes
Cow	Yes	No	No
Eagle	No	Yes	Yes
Dog	Yes	No	Yes

Show the steps performed by the ID3 algorithm in building a decision tree that classifies these examples. That is, show the calculations used to determine which properties to split on, in the appropriate order.

Human vs Machine Intelligence

- The two hemispheres of the human brain deal with problems in two distinct paradigms:
  - sequential (or logical) approach that considers only a small portion of the available data at a time
  - parallel processing looks at data on a global basis

Many tasks which we might reasonably think require **intelligence** are performed by computers **without even thinking** → Complex Arithmetic

Other tasks that people do without thinking are extremely difficult to automate → Recognizing a Face

Symbolic vs. sub-symbolic AI

"Old"-Fashioned AI is inherently symbolic

Physical Symbol System Hypothesis: ***A necessary and sufficient condition for intelligence is the representation and manipulation of symbols.***

alternatives to symbolic AI

A

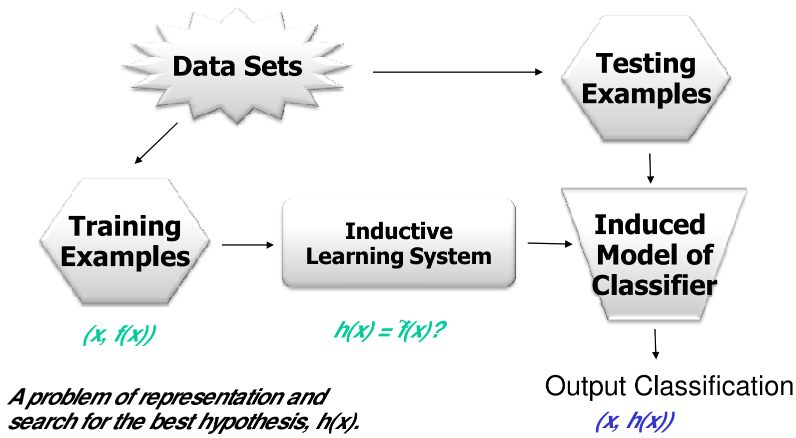
- connectionist models – based on a brain metaphor
  model individual neurons and their connections
  properties: parallel, distributed, sub-symbolic
  examples: neural nets, perceptrons, backpropagation NN, associative memories (Hopfield networks)

SubSymbolic Representations

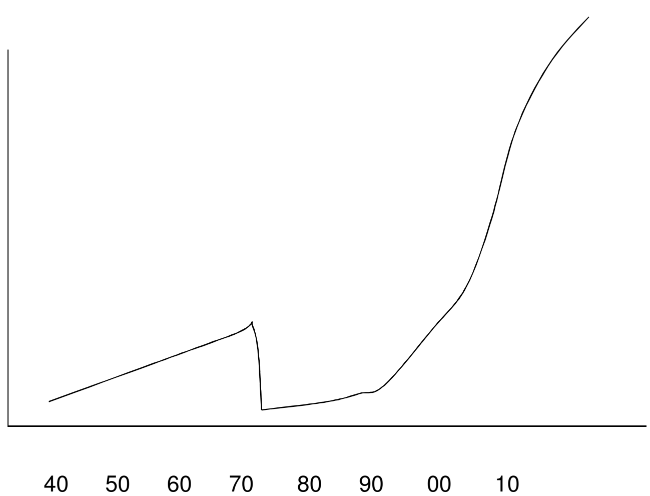
Decision trees can be easily read
  A disjunction of conjunctions (logic), We call this a symbolic representation
Non-symbolic representations
  More numerical in nature, more difficult to read
Artificial Neural Networks (ANNs)
  A Non-symbolic representation scheme
  They embed a giant mathematical function
    To take inputs and compute an output which is interpreted as a categorisation
  Often shortened to "Neural Networks"
    Don't confuse them with real neural networks (in heads)

Classification Systems and Inductive Learning

Basic Framework for Inductive Learning

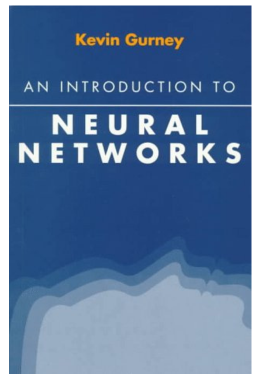


Interest in Subsymbolic AI



Helpful Resource (in addition to Luger)

Gurney, Kevin. *An Introduction to Neural Networks*, 1996.



Neural Networks

- Neural Networks can be :
    - **Biological** models
    - **Artificial** models
- Desire to produce artificial systems capable of sophisticated computations similar to the human brain.

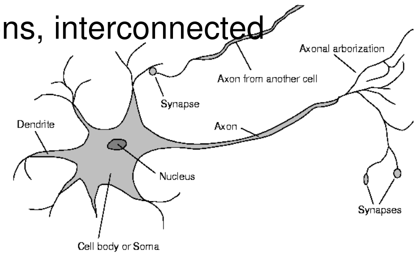
Biological analogy

general brain architecture:

- many (relatively) slow neurons, interconnected

Dendrites(التغصنات)

- (the receivers) serve as input devices (receive electrical impulses from other neurons)



- axon (the transmitter synapse) point of transmission
- cell body "sums" inputs from the dendrites (possibly inhibiting or exciting)
- if sum exceeds some threshold, the neuron fires an output impulse along axon, sending an electrical pulse to other neurons

Biological analogy

The brain is composed of a mass of interconnected neurons, each neuron is connected to many other neurons

The brain:

10^{11} neurons of > 20 types, 10^{14} synapses, 1ms–10ms cycle time

Signals are noisy "spike trains" of electrical potential

Neurons transmit signals to each other

Whether a signal is transmitted to an event or not (the electrical potential in the cell body of the neuron is thresholded)

Whether a signal is sent, depends on the strength of the bond (synapse-المشبك) between two neurons

From Biological to Artificial Neurons

An Artificial Neuron - The Perceptron

simulated on hardware or by software

input connections - the receivers

node, unit, or *PE* simulates neuron body

output connection - the transmitter

activation function employs a threshold or *bias*

connection weights act as synaptic junctions

Learning occurs via changes in value of the connection weights.

From Biological to Artificial Neurons

An Artificial Neuron - The Perceptron

Basic function of neuron is to sum inputs, and produce output given sum is greater than threshold

ANN node produces an output as follows:

1. Multiplies each component of the input pattern by the weight of its connection
2. Sums all weighted inputs and subtracts the threshold value => *total weighted input*
3. Transforms the total weighted input into the output using the activation function

Connectionist models (neural nets)

Definitions

Neuron: The cell that performs information processing in the brain.

- Fundamental functional unit of all nervous system tissue.

A Neural Network: is a system composed of many simple processing elements operating in parallel which can acquire, store, and utilize experiential knowledge.

- Each element of NN is a node called unit.
- Units are connected by links.
- Each link has a numeric weight.

History of Neural Networks

1943: McCulloch and Pitts proposed a model of a neuron --> Perceptron

1960: Widrow and Hoff explored Perceptron networks (which they called "Adelines")

1962: Rosenblatt proved the convergence of the perceptron training rule.

1969: Minsky and Papert showed that the Perceptron cannot deal with nonlinearly-separable data sets---even those that represent simple function such as X-OR.

1970-1985: Very little research on Neural Nets

1986: Invention of Backpropagation which can learn from nonlinearly-separable data sets. [Rumelhart and McClelland, but also Parker and earlier on: Werbos]

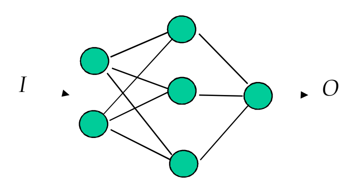
Since 1985: A lot of research in Neural Nets!

ANN

Inherent Advantages of the Brain:

"distributed processing and representation"

- Parallel processing speeds
- Fault tolerance
- Ability to generalize



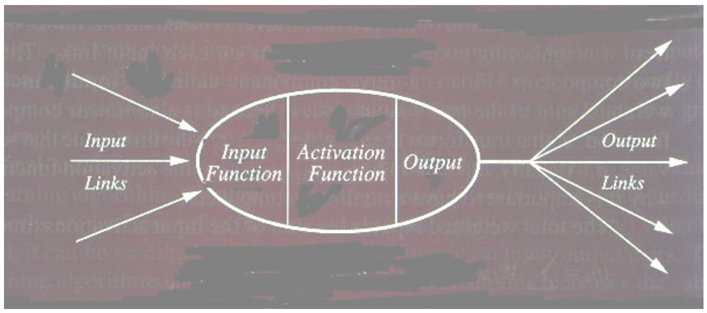
How NN learns a task.

Issues to be discussed

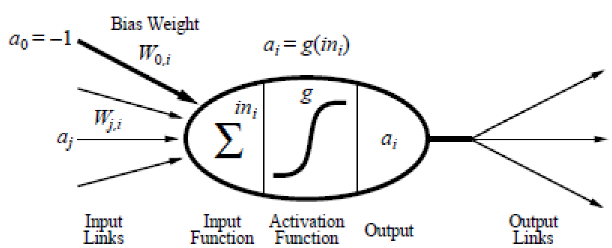
- Initializing the weights.
- Use of a learning algorithm.
- Set of training examples.
- Encode the examples as inputs.
- Convert output into meaningful results.

Computing Elements

A typical unit:



Computing Elements



The output : $a_i = g(\Sigma_j W_{j,i} a_j)$

Simple Computations in this network

-There are **2 types of components:** Linear and Non-linear.

- **Linear:** Input function
  - calculate weighted sum of all inputs.

- **Non-linear:** Activation function
  - transform sum into activation level.

Activation Functions

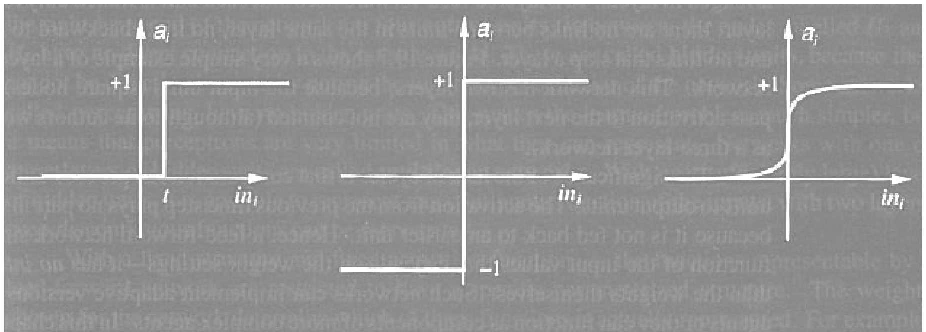
- Use different functions to obtain different models.

- 3 most common choices :

  1) Step function
  2) Sign function
  3) Sigmoid function

- An output of **1 represents firing** of a neuron down the axon.

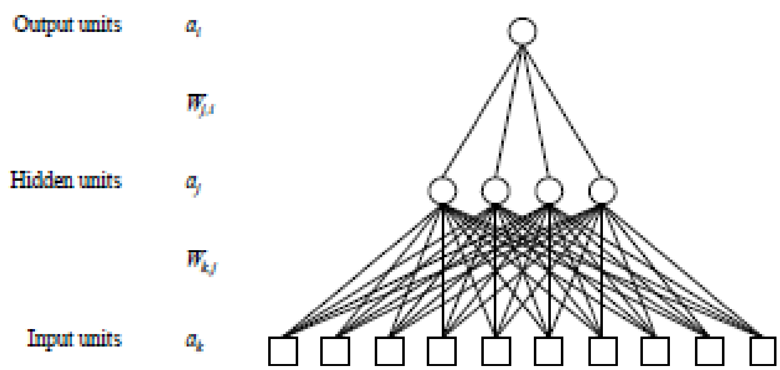
Activation Functions



(a) Step function (b) Sign function (c) Sigmoid function

$$\text{step}_t(x) = \begin{cases} 1, & \text{if } x \geq t \\ 0, & \text{if } x < t \end{cases} \qquad \text{sign}(x) = \begin{cases} +1, & \text{if } x \geq 0 \\ -1, & \text{if } x < 0 \end{cases} \qquad \text{sigmoid}(x) = \frac{1}{1 + e^{-x}}$$

Two-Layer Feed-Forward Neural Network



Standard structure of an artificial neural network

Input units

represents the input as a fixed-length vector of numbers (user defined)

Hidden units

calculate threshold weighted sums of the inputs

represent intermediate calculations that the network learns

Output units

represent the output as a fixed length vector of numbers

Network Structure

Feed-forward neural nets:

Links can only go in one direction.

Recurrent neural nets:

Links can go anywhere and form arbitrary topologies.

Feed Forward Network

- Arranged in ***layers***.
- Each unit is linked only in the unit in next layer.
- No units are linked between the same layer, back to the previous layer or skipping a layer.

- Computations can proceed uniformly from input to output units.

Multi-layer Networks

- Have one or more layers of hidden units.

- With two possibly very large hidden layers, it is possible to implement any function

Perceptrons

- Networks without hidden layer are called perceptrons.

- Perceptrons are very limited in what they can represent, but this makes their learning problem much simpler.