ARTIFICIAL NEURAL NETWORKS: A TUTORIAL

BY:

Negin Yousefpour

PhD Student

Civil Engineering Department

TEXAS A&M UNIVERSITY

Contents

- Introduction
- Origin Of Neural Network
- Biological Neural Networks
- ANN Overview
- Learning
- Different NN Networks
- Challenging Problems
- Summery

INTRODUCTION

- Artificial Neural Network (ANN) or Neural Network(NN) has provide an exciting alternative method for solving a variety of problems in different fields of science and engineering.
- This article is trying to give the readers a :

- Whole idea about ANN
- Motivation for ANN development
- Network architecture and learning models
- Outline some of the important use of ANN

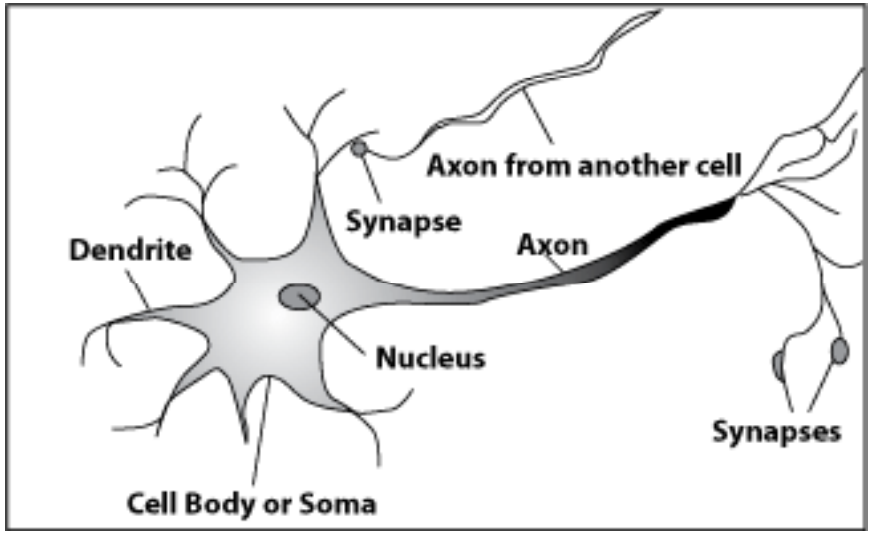
Origin of Neural Network

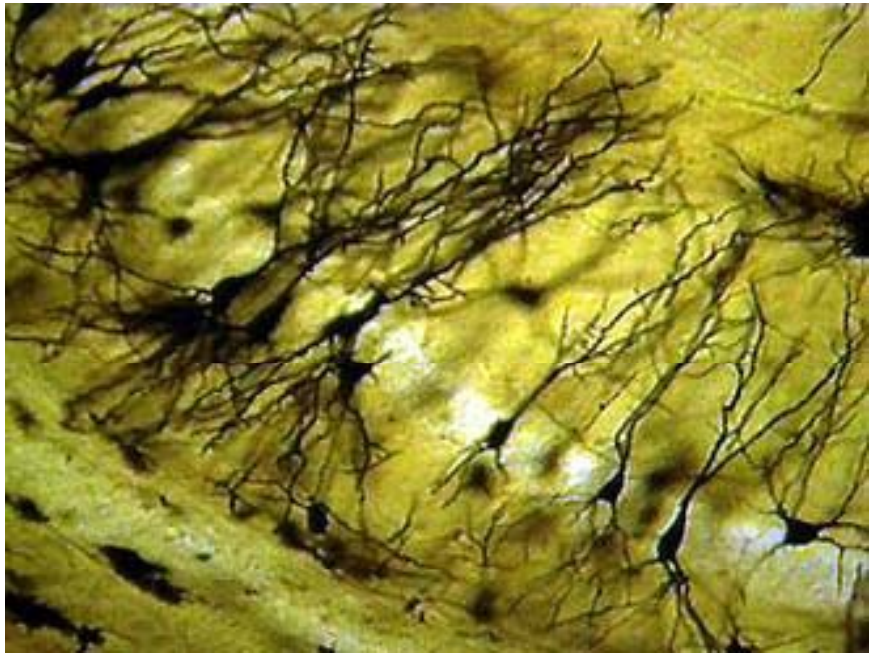
- Human brain has many incredible characteristics such as massive parallelism, distributed representation and computation, learning ability, generalization ability, adaptivity, which seems simple but is really complicated.
- It has been always a dream for computer scientist to create a computer which could solve complex perceptual problems this fast.
- ANN models was an effort to apply the same method as human brain uses to solve perceptual problems.
- Three periods of development for ANN:

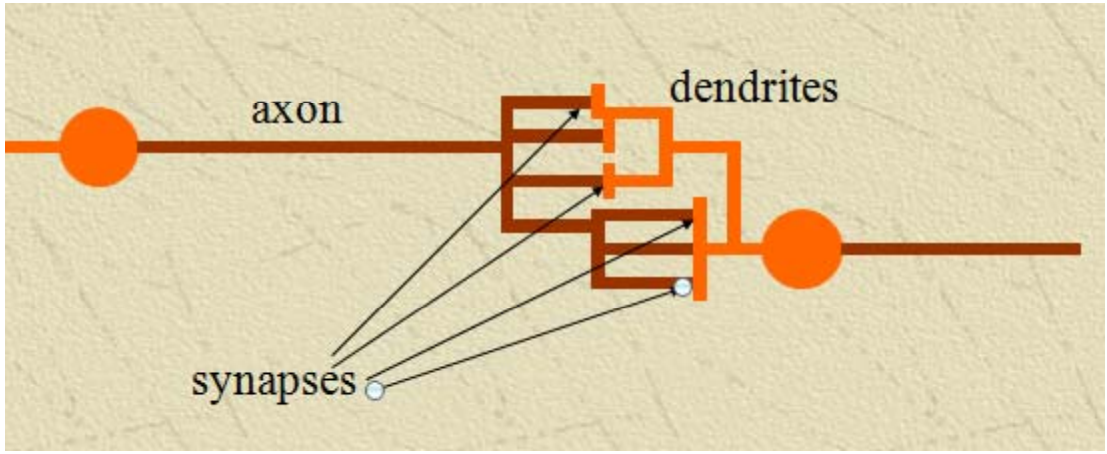
- **1940**:Mcculloch and Pitts: Initial works
- **1960**: Rosenblatt: perceptron convergence theorem

  Minsky and Papert: work showing the limitations of a simple perceptron
- **1980**: Hopfield/Werbos and Rumelhart: Hopfield's energy approach/back-propagation learning algorithm

Biological Neural Network

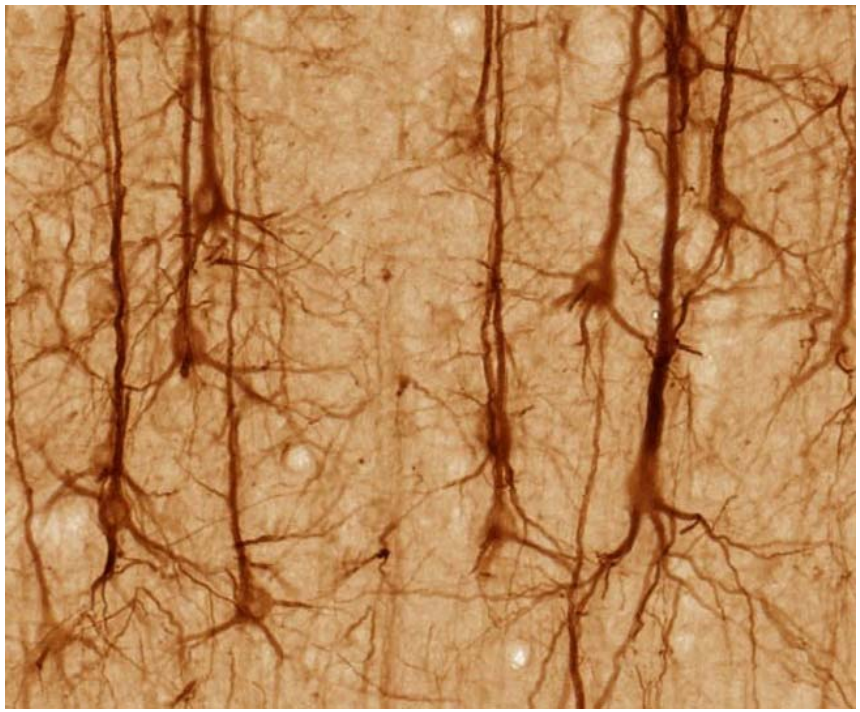




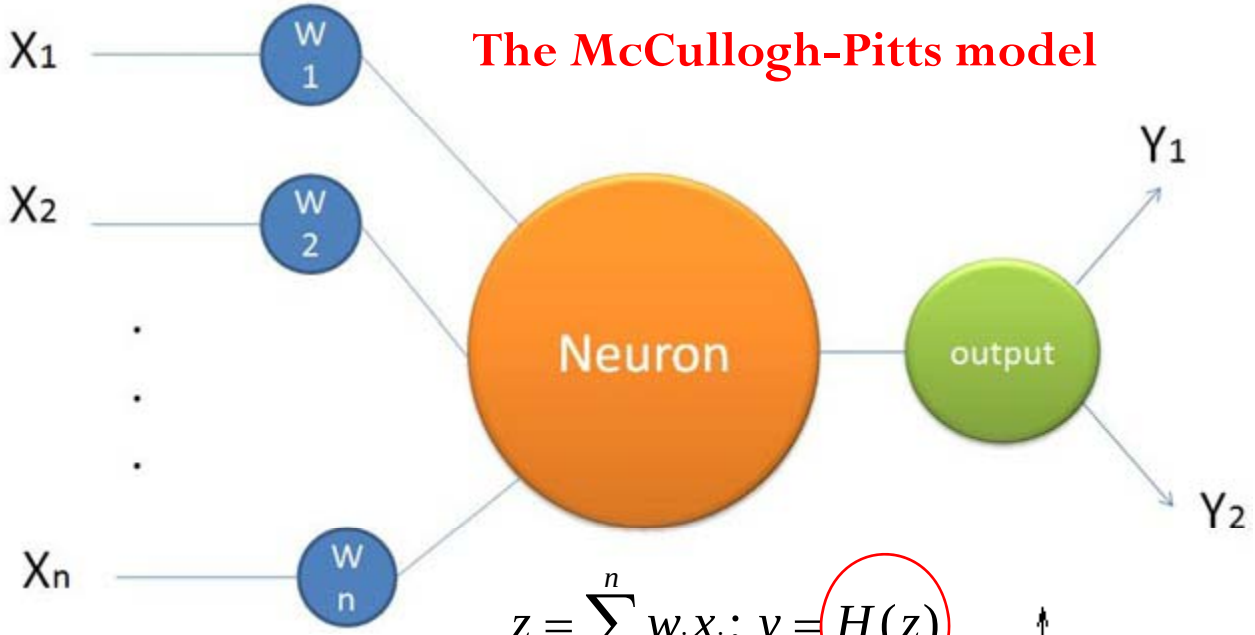


Biological Neural Network

- When a signal reaches a synapse: Certain chemicals called neurotransmitters are released.
- Process of learning: The synapse effectiveness can be adjusted by signal passing through.
- Cerebral cortex :a large flat sheet of neurons about 2 to 3 mm thick and 2200 cm , 10^11 neurons
- Duration of impulses between neurons: milliseconds and the amount of information sent is also small(few bits)
- Critical information are not transmitted directly , but stored in interconnections
- The term Connectionist model initiated from this idea.

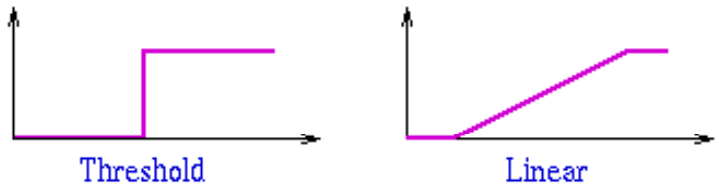


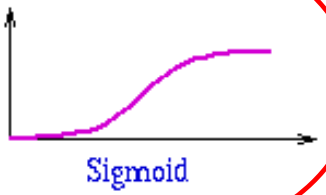
ANN Overview: COMPTIONAL MODEL FOR ARTIFICIAL NEURON



The McCullogh-Pitts model

$$z = \sum_{i=1}^{n} w_i x_i ; \; y = H(z)$$



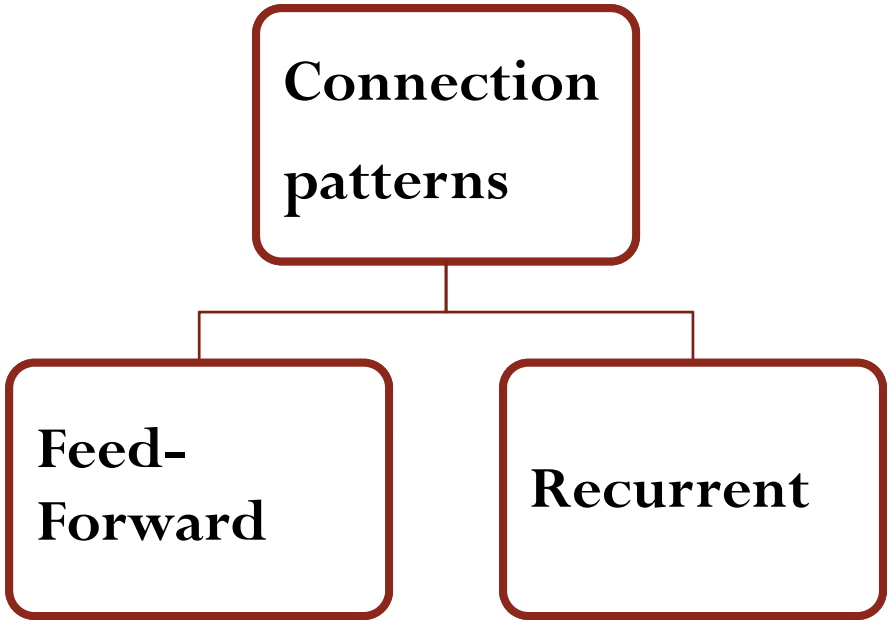


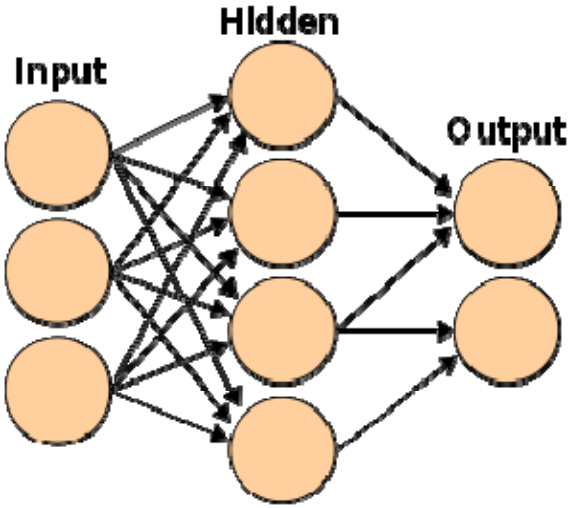
Wires : axon & dendrites

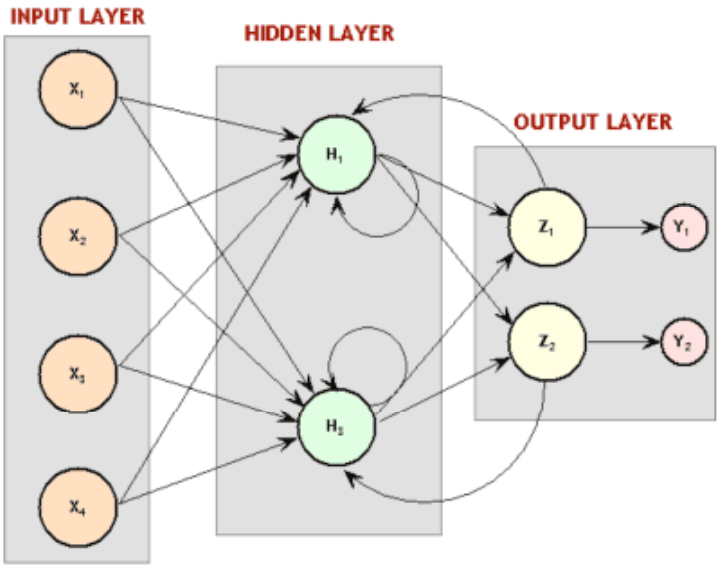
Connection weights: Synapse

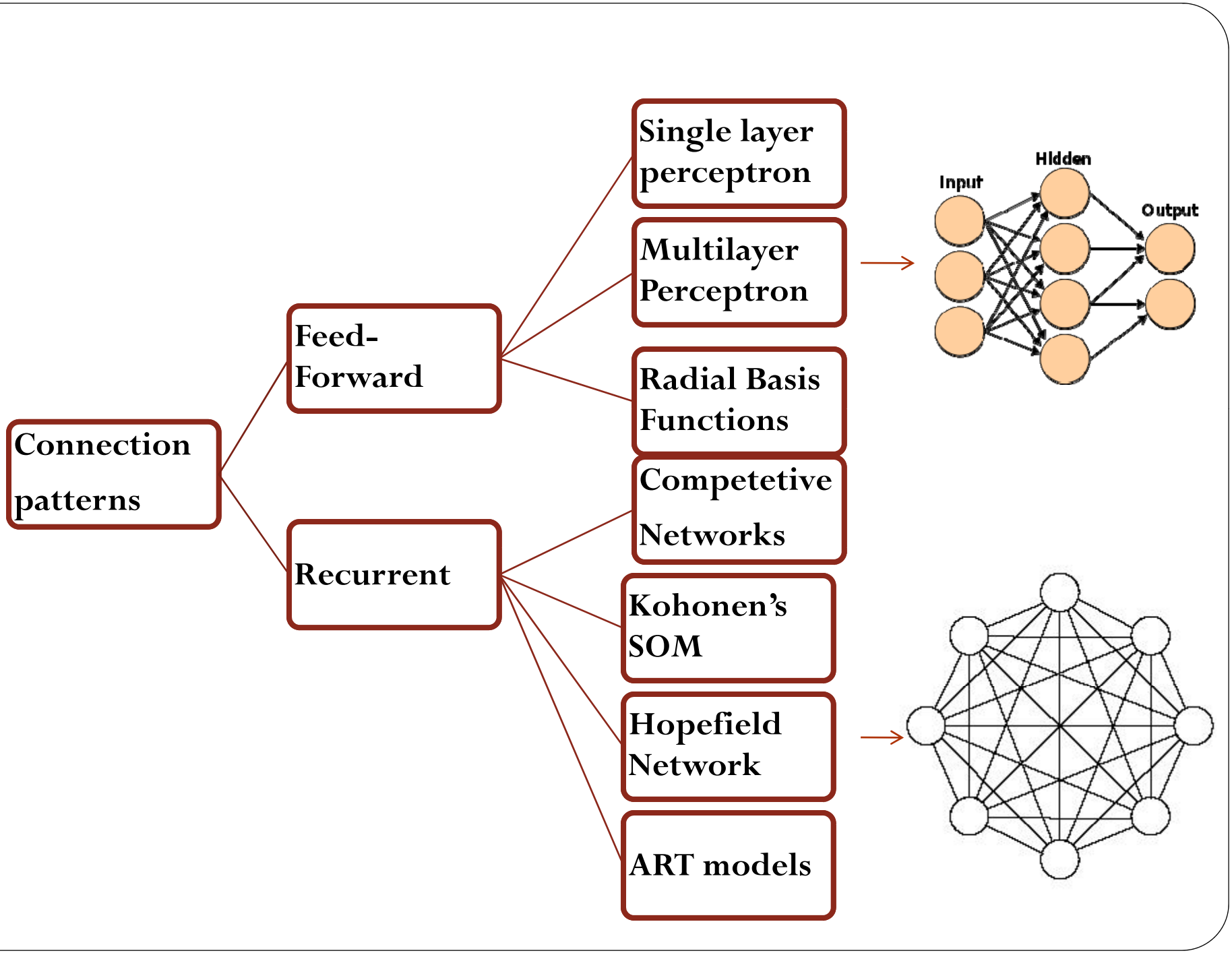
Threshold function: activity in soma

ANN Overview: Network Architecture








Connection patterns
Feed-Forward
Recurrent
Single layer perceptron
Multilayer Perceptron
Radial Basis Functions
Competetive Networks
Kohonen's SOM
Hopefield Network
ART models
Input
Hidden
Output

Learning

- **What is the learning process in ANN?**
- updating network architecture and connection weights so that network can efficiently perform a task
- **What is the source of learning for ANN?**
- Available training patterns
- The ability of ANN to automatically learn from examples or input-out put relations
- **How to design a Learning process?**
- Knowing about available information
- Having a model from environment: Learning Paradigm
- Figuring out the update process of weights: Learning rules
- Identifying a procedure to adjust weights by learning rules: Learning algorithm

Learning Paradigm

1. Supervised

- The correct answer is provided for the network for every input pattern
- Weights are adjusted regarding the correct answer
- In reinforcement learning only a critique of correct answer is provided

2. Unsupervised

- Does not need the correct output
- The system itself recognize the correlation and organize patterns into categories accordingly

3. Hybrid

- A combination of supervised and unsupervised
- Some of the weights are provided with correct output while the others are automatically corrected.

Learning Rules

- There are four basic types of learning rules:

- ***Error correction rules***

- ***Boltzmann***

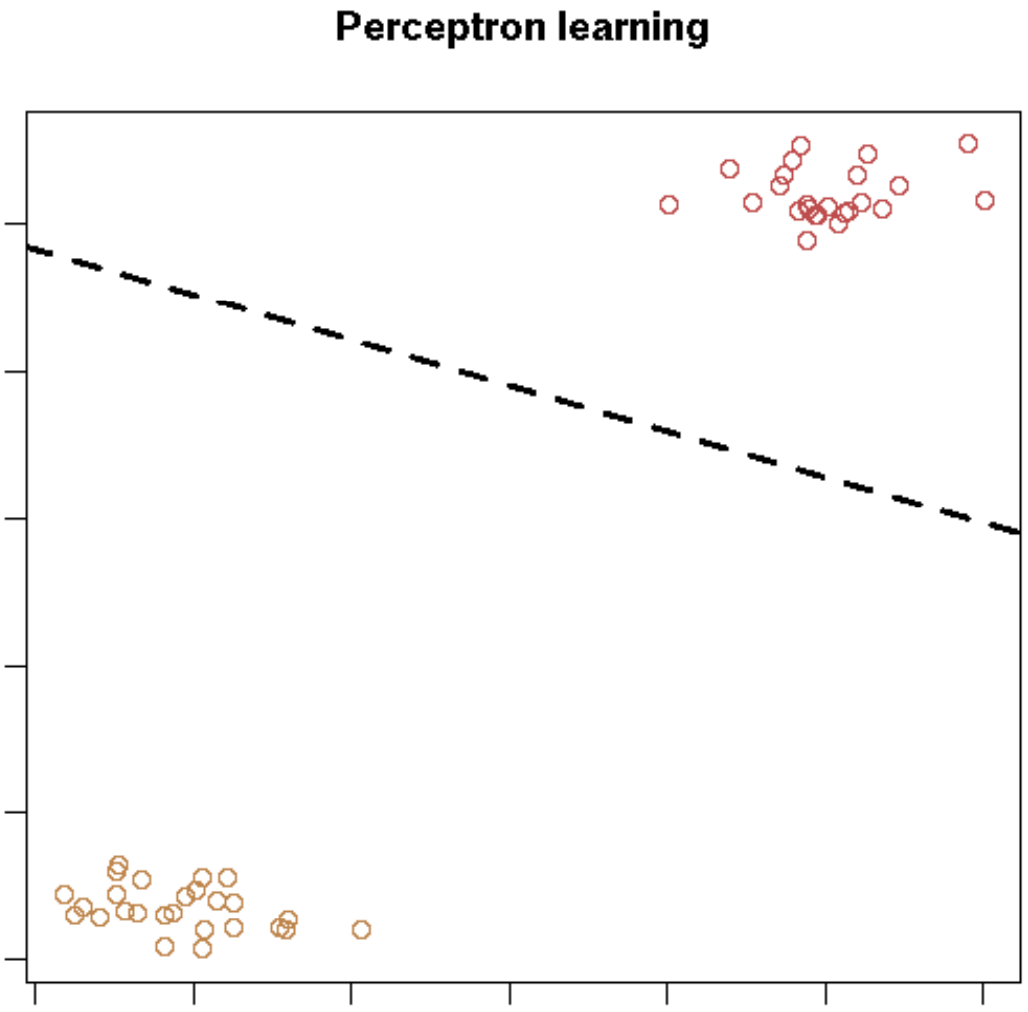
- ***Hebbian***

- ***Competitive learning***

- Each of these can be trained with or with out a teacher
- Have a particular architecture and learning algorithm

Error Correction Rules

- An e ed to modify the c
- Perc or-correction princ
- A pe th adjustable weig
- If an iterating unit ill re
- Since patterns are linea ith an infinite No. c

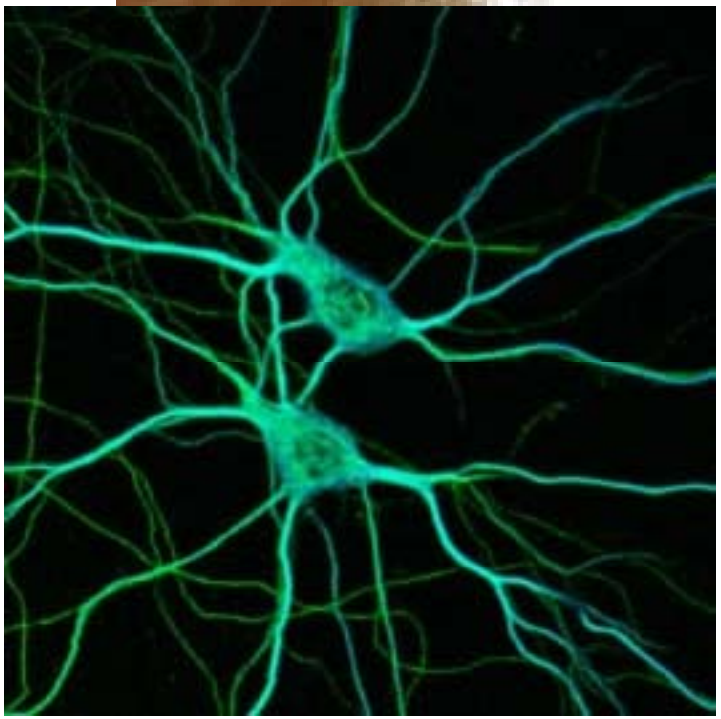


Boltzman Learning

- Used in symmetric recurrent networks(symmetric:Wij=Wji)
- Consist of binary units(+1 for on, -1 for off)
- Neurons are divided into two groups: Hidden & Visible
- Outputs are produced according to Boltzman statistical mechanics
- Boltzman learning adjust weights until visible units satisfy a desired probabilistic distribution
- The change in connection weight or the error-correction is measured between the correlation between two pair of input and output neuron under clamped and free-operating condition

Hebbian Rules

- One of the oldest learning rule initiated form neurobiological experiments
- The basic concept of Hebbian Learning: when neuron A activates, and then causes neuron B to activate, then the connection strength between the two neurons is increased, and it will be easier for A to activate B in the future.
- Learning is done locally, it means weight of a connection is corrected only with respect to neurons connected to it.
- Orientation selectivity: occurs due to Hebbian training of a network

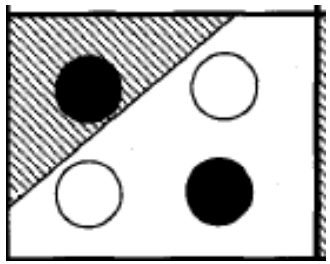
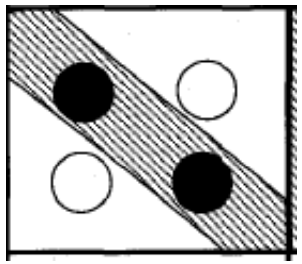
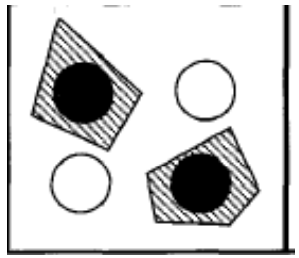


Competitive Learning Rules

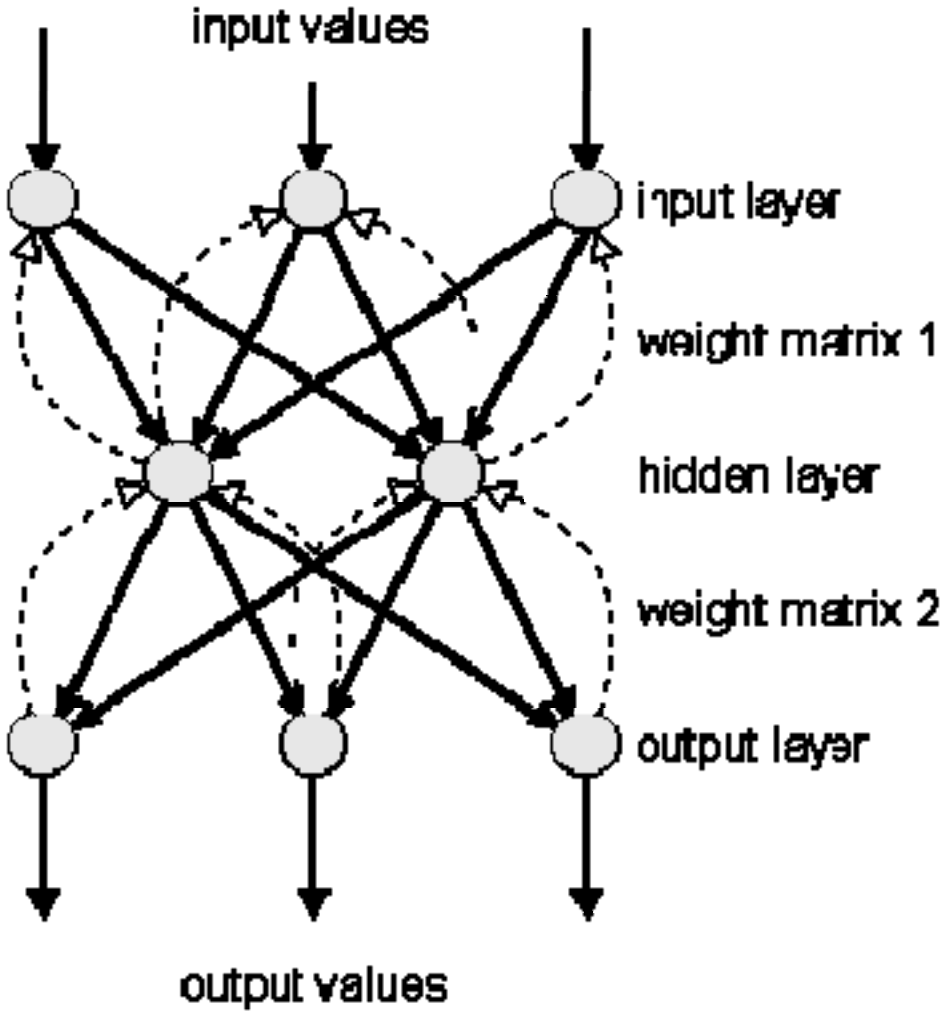
- The basis is the "winner take all" originated from biological Neural network
- All input units are connected together and all units of output are also connected via inhibitory weights but gets feed back with excitory weight
- Only one of the unites with largest or smallest input is activated and its weight becomes adjusted
- As a result of learning process the pattern in the winner unit (weight) become closer to he input pattern.
- http://www.peltarion.com/blog/img/sog/competitive.gif

Multilayer Perceptron

- The most popular networks with feed-forward system
- Applies the Back Propagation algorithm
- As a result of having hidden units, Multilayer perceptron can form arbitrarily complex decision boundaries
- Each unit in the first hidden layer impose a hyperplane in the space pattern
- Each unit in the second hidden layer impose a hyperregion on outputs of the first layer
- Output layer combines the hyperregions of all units in second layer

Back propagation Algorithm



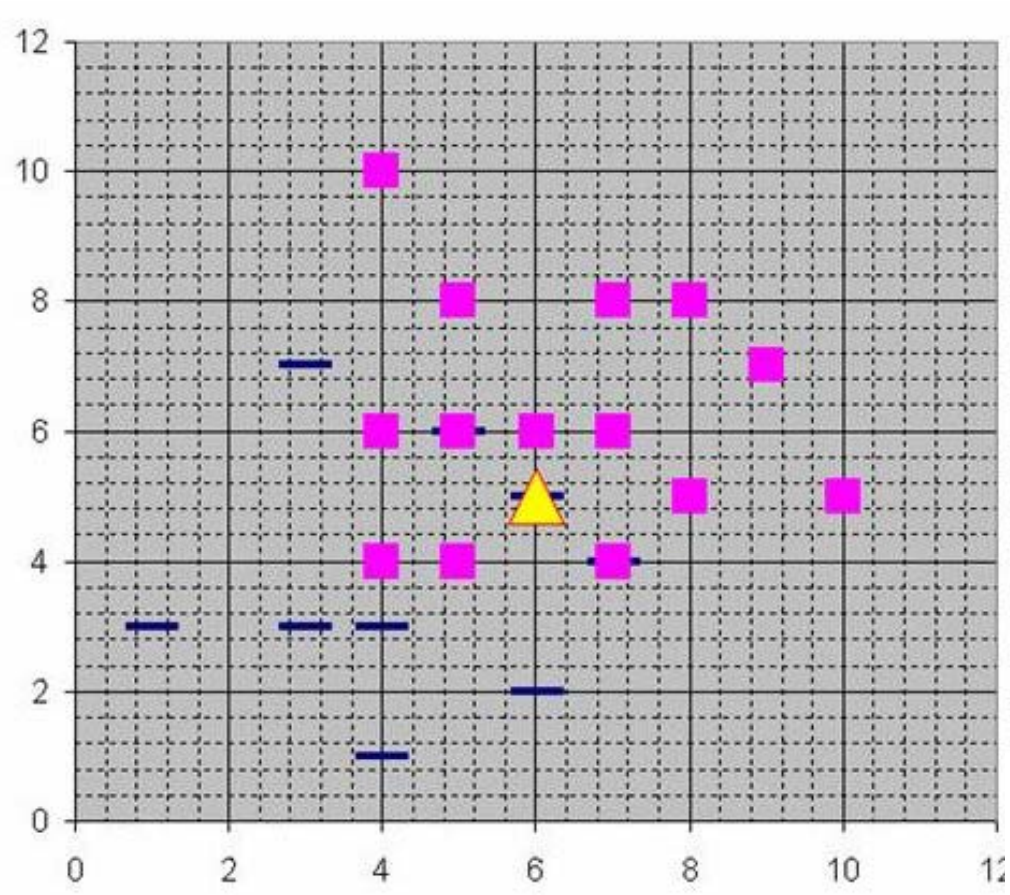
Radial Basis Function Network(RBF)

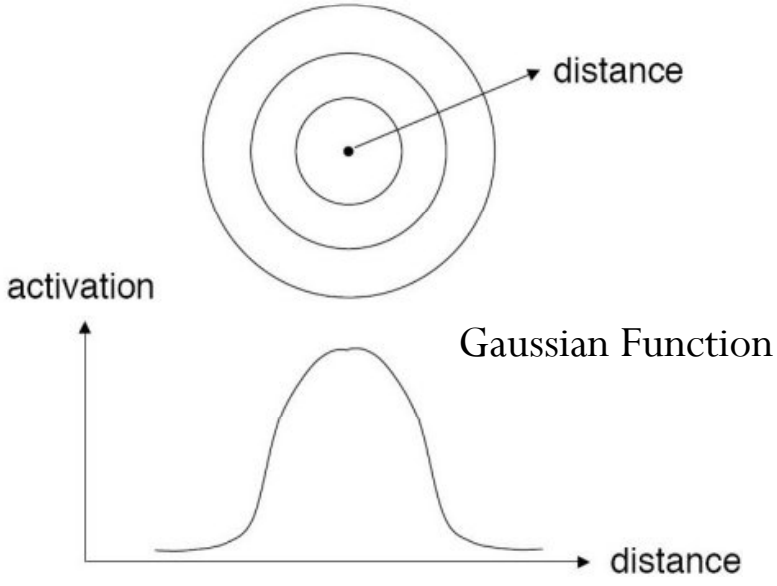
- A Radial Basis Function like Gaussian Kernel is applied as an activation function.
- A Radial Basis Function(also called Kernel Function) is a real-valued function whose value depends only on the distance from the origin or any other center: $F(x)=F(|x|)$
- RBF network uses a hybrid learning , unsupervised clustering algorithm and a supervised least square algorithm
- As a comparison to multilayer perceptron Net.:

  -The Learning algorithm is faster than back-propegation

  - After training the running time is much more slower

Let's see an example!

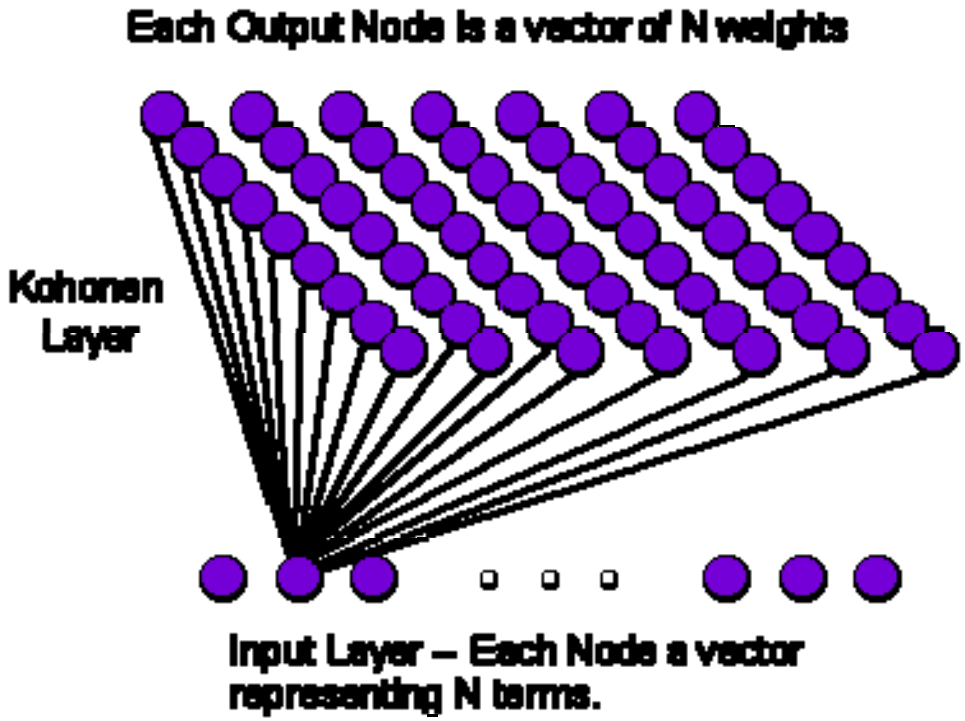




Gaussian Function

Kohonen Self-Organizing Maps

- It consists of a two dimensional array of output units connected to all input nodes
- It works based on the property of Topology preservation
- Nearby input patterns should activate nearby output units on the map
- SOM can be recognized as a special competitive learning
- Only the weight vectors of winner and its neighbor units are updated

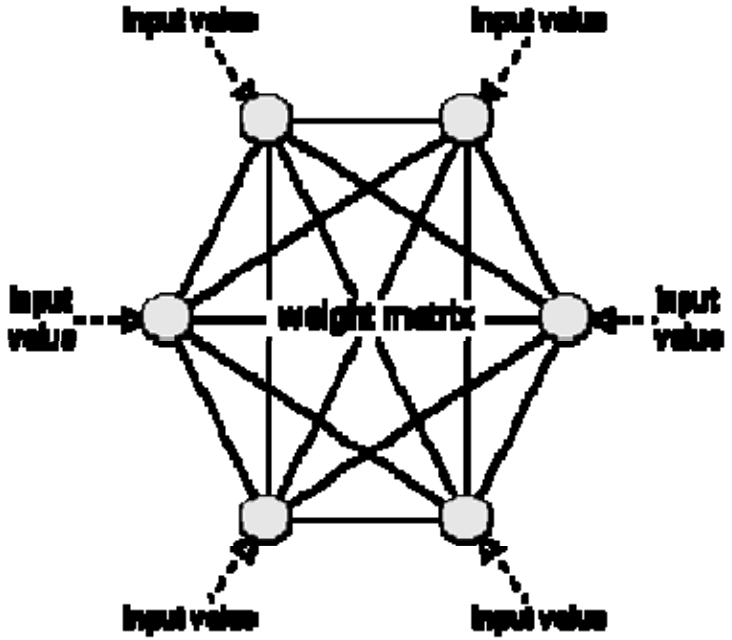


Adaptive Resonance Theory Model

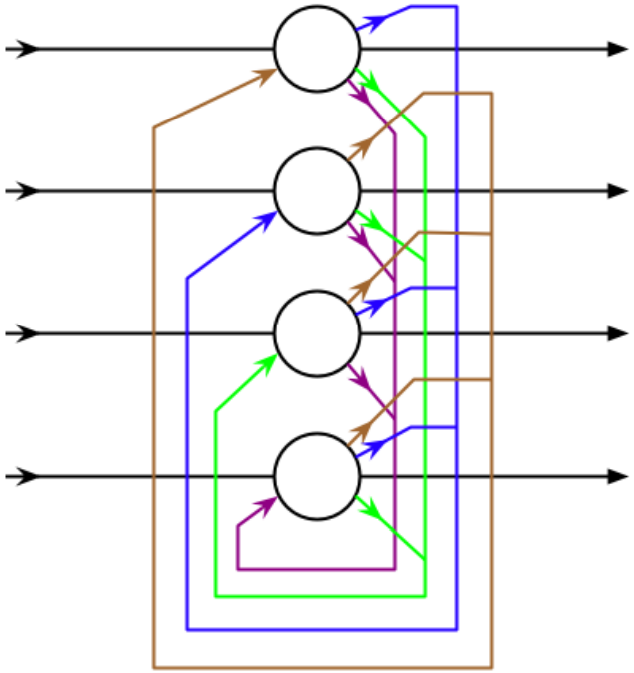
- ART models were proposed theories to overcome the concerns related to stability-plasticity dilemma in competitive learning.
- Is it likely that learning could corrupt the existing knowledge in a unite?
- If the input vector is similar enough to one of the stored prototypes(resonance), learning updates the prototype.
- If not, a new category is defined in an "uncommitted" unit.
- Similarity is controlled by vigilance parameter.

Hopfield Network

- Hopfield designed a network based on an energy function
- As a result of dynamic recurrency, the network total energy decreases and tends to a minimum value(attractor)
- The dynamic updates take places in two ways: synchronously and asynchronously
- traveling salesman problem(TSP)

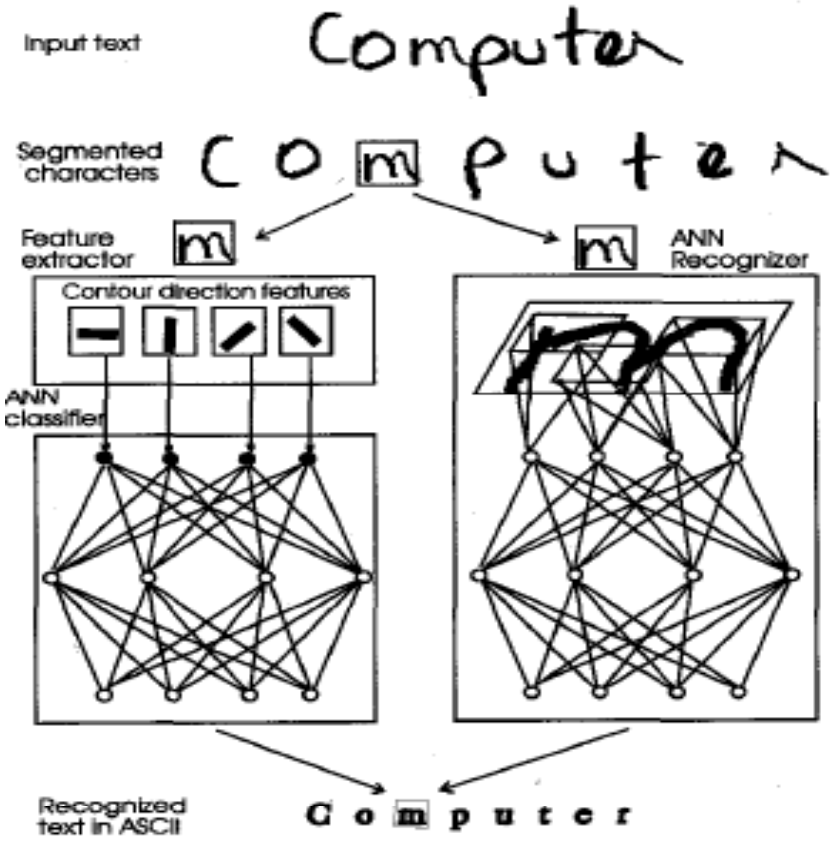


Hopfield Net

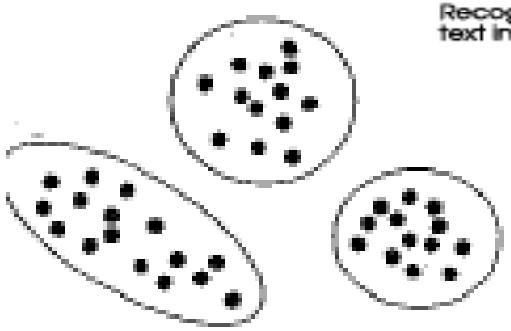


Challenging Problems

- Pattern recognition
  - Character recognition
  - Speech recognition



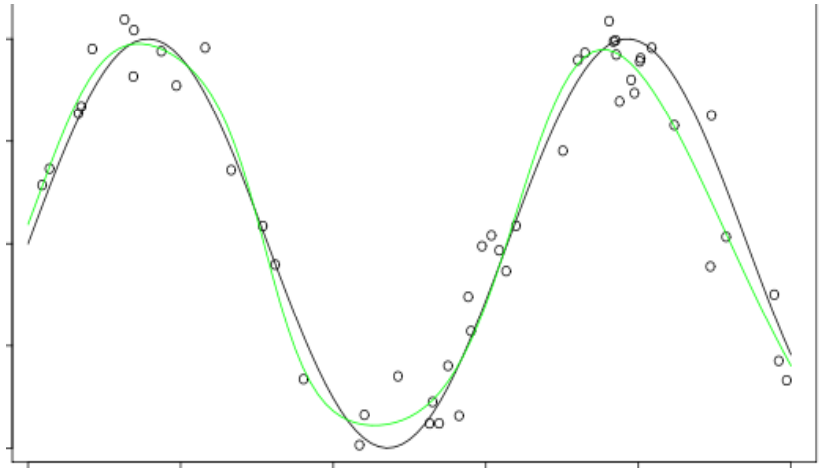
- Clustering/Categorization
  - Data mining
  - Data analysis



Challenging Problems

- Function approximation

  - Engineering & scientific Modeling



- Prediction/Forecasting

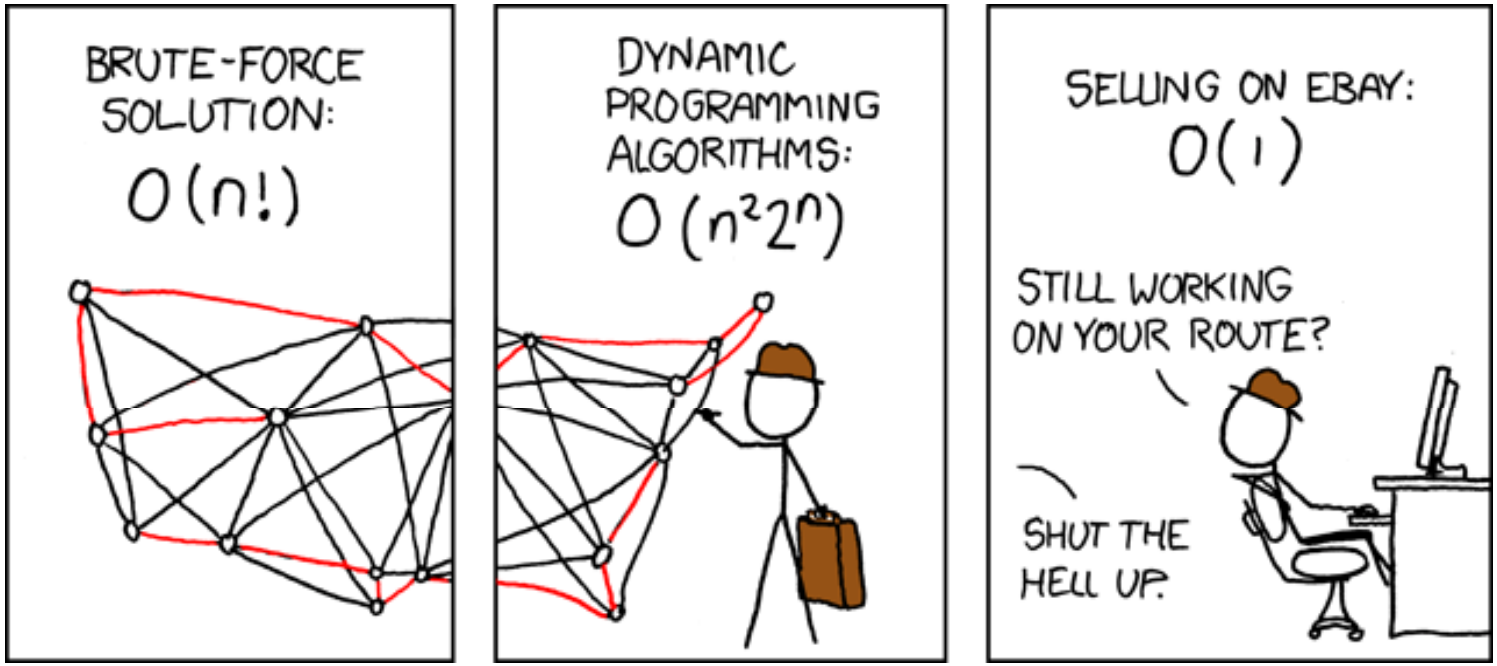
  - Decision-making

  - Weather forecasting

Challenging Problems

- Optimization

- Traveling salesman problem(TSP)



Summery

- A great overview of ANN is presented in this paper, it is very easy understanding and straightforward
- The different types of learning rules, algorithms and also different architectures are well explained
- A number of Networks were described through simple words
- The popular applications of NN were illustrated
- The author Believes that ANNS brought up both enthusiasm and criticism.
- Actually except for some special problems there is no evidence that NN is better working than other alternatives
- More development and better performance in NN requires the combination of ANN with new technologies