In The Name of God
Introduction to Artificial Neural Networks

Lecture 1:

Introduction

By: Ali Motie Nasrabadi
Assistant professor
Faculty of Engineering
University of Shahed

Outline

- **Introducing the course**
  - **Aims and Learning Outcomes**
  - **Assessment**
  - **Lecture Plan**
  - **Recommended Books**
  - **Comments on Mathematical Requirements**
- **(Artificial) Neural Network ?**
- **Basic concepts**
- **Neural Network Definitions**
- **Basic characteristic of biological neurons**
- **Benefits of Neural Networks**
- **Who is concerned with NNs?**
- **Some Current Artificial Neural Network Applications**

Aims and Learning Outcomes

■Aims

- **1. Introduce some of the fundamental techniques and principles of neural network systems.**
- **2. Investigate some common models and their applications.**

■Learning Outcomes

- **1. Understand the relation between real brains and simple artificial neural network models.**
- **2. Describe and explain the most common architectures and learning algorithms for Multi-Layer Perceptrons, Radial-Basis Function Networks and Kohonen Self-Organizing Maps (SOM), Hopfield , Adaptive Resonance Theory (ART), …**
- **3. Explain the learning and generalization aspects of neural network systems.**
- **4. Demonstrate an understanding of the implementational issues for common neural**
- **5. Demonstrate an understanding of the practical considerations in applying neural networks to real classification, recognition and approximation problems.**

Assessment

- **50% Closed book examination**
- **20% Homework**
- **20% Final project**
- **10% Class Interactive Participation**

Lecture Plan(1)

- Introduction to Neural Networks and their History. Biological &Neurons and Neural Networks. Artificial Neurons.
- Pattern recognition and regression
- Networks of Artificial Neurons. Single Layer Perceptrons Learning and Generalization in Single Layer Perceptrons
- Basic structures and properties of Artificial Neural Networks
- Introduction to learning: Hebbian Learning. Gradient Descent Learning. The Generalized Delta Rule. Practical Considerations.
- Learning in Multi-Layer Perceptrons. Back-Propagation Learning with Momentum. Conjugate Gradient Learning

Lecture Plan(2)

- Bias and Variance. Under-Fitting and Over-Fitting. Improving Generalization.
- Applications of Multi-Layer Perceptrons.
- Radial Basis Function Networks: Introduction, Algorithms and Applications.
- Competitive Learning and Self Organizing Maps: Fundamentals, Algorithms and Applications.
- Learning Vector Quantization (LVQ).
- Hopfield net: Fundamentals, Algorithms and Applications
- Adaptive Resonance Theory (ART): Fundamentals, Algorithms and Applications

Recommended Books

- **Haykin, S.**
  ***Neural Networks. A Comprehensive Foundation.*,**
  Second Edition, Prentice-Hall, Inc., New Jersey, 1999.
- **R. Beale & T. Jackson**
  Neural Computing : An *Introduction*
  IOP Publishing, 1990
- **Hagan,M.T., Demuth,H.B. and Beale,M.H.**
  ***Neural Network Design*,**
  PWS Publishing Co., Boston, MA, 1996.
- **Hertz,J., Krogh,A., and and Palmer,R.G.**
  ***Introduction to The Theory of Neural Computation*,**
  Addison-Wesley Publishing Company Inc., Reading, MA, 1991.
- **Zurada,J.M.**
  ***Introduction to Artificial Neural Systems*,**
  PWS Publishing Co., 1992.
- **Bishop, C.**
  Neural Networks for Pattern Recognition
  Oxford University Press, 1995.

Table of Ref.

Main Recommended Books

Title	Author(s)	Publisher, Date	Comments
Neural Networks: A Comprehensive Foundation	Simon Haykin	Prentice Hall, 1999	Very comprehensive and up-to-date, but heavy in maths.
An Introduction to Neural Networks	Kevin Gurney	UCL Press, 1997	Non-mathematical introduction.
Neural Networks for Pattern Recognition	Christopher Bishop	Clarendon Press, Oxford, 1995	This is the book I always use.
The Essence of Neural Networks	Robrt Callan	Prentice Hall Europe, 1999	Concise introductory text.

Other Good Books

Title	Author(s)	Publisher, Date	Comments
Introduction to Neural Networks	R. Beale & T. Jackson	IOP Publishing, 1990	Former recommended book.
An Introduction to the Theory of Neural Computation	J. Hertz, A. Krogh & R.G. Palmer	Addison Wesley, 1991	Good all round book. Slightly mathematical.
Parallel Distributed Processing: Volumes 1 and 2	D.E. Rummelhart, J.L. McClelland, et al.	MIT Press, 1986	The original neural networks bible.
The Computational Brain	P.S. Churchland & T.J. Sejnowski	MIT Press, 1994	Good for computational neuroscience.
Principles of Neurocomputing for Science and Engineering	F.M. Ham & I. Kostanic	McGraw Hill, 2001	Good advanced book, but rather mathematical.

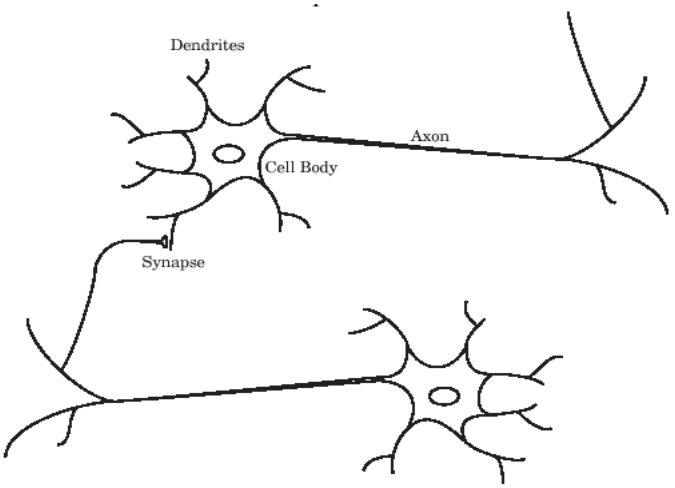
Reference of Presentation

- M. Shahram Moin, "Artificial Neural Networks", Iran telecommunication research center
- Jin Hyung Kim, "Neural Networks: Introduction", Computer Science Department KAIST
- Jianfeng Feng, "Introduction to Neural Networks", School of Cognitive and Computing Sciences, jianfeng@cogs.susx.ac.uk
- John A. Bullinaria, "Introduction to Neural Networks http://www.cs.bham.ac.uk/~jxb/inn.html
- Jonathan Shapiro, "Neural Networks", Department of Computer Science, University of Manchester, jls@cs.man.ac.uk.
- Howard Demuth, Mark Beale, "Neural Network Toolbox"
- Paul Sajda," Computational Neural Modeling and Neuroengineering", http://www.bme.columbia.edu/~sajda/bme6480
- Ricardo Gutierrez-Osuna," Introduction to Pattern Analysis", Texas A&M University, http://faculty.cs.tamu.edu/rgutier/
- William H. Hsu, "A Brief Survey of Machine Learning", Department of Computing and Information Sciences, KSU, http://www.cis.ksu.edu/~bhsu

Comments on Mathematical Requirements

- The easiest way to formulate and understand neural networks is in terms of mathematical concepts and equations (Vector and Matrix).
- Once you have the equations it is fairly straightforward to convert them into C/C++/Pascal/MATLAB programs.
- This module will introduce and explain any necessary mathematics as and when we need it. Much of this will also be useful for other modules, such as Machine Learning.
- You will not be required to perform any mathematical derivations in the examination.

The Neuron



(Artificial) Neural Network ?

- **Computational model inspired from neurological model of brain**
- **Human brain computes in different way from digital computer**
  - **highly complex, nonlinear, and parallel computing**
  - **many times faster than d-computer in**
    - **pattern recognition, perception, motor control**
  - **has great structure and ability to build up its own rules by experience**
    - **dramatic development within 2 years after birth**
    - **continues to develop afterward**
  - **Plasticity : ability to adapt to its environment**

Basic concepts (1)

- **No universally accepted definition**
- **Most people in the field agree that:**
  - **An NN is a network of many *simple processors (units)***
  - **Each processor has a small amount of *memory***
  - **Units are connected by communication channels (*connections*)**
- **Some NNs are models of biological neural networks and some are not**
- **Historically, inspiration came from desire to produce artificial systems capable of *sophisticated,* perhaps *intelligent* computations similar to those that the human "*brain"* performs**
- **Enhance our understanding of the human brain**

Basic concepts (2)

- **Training**
  - **A set of rules**
  - **Adaptation,**
  - **W*eights* of connections are *adjusted* on the basis of data**
- **Generalization**
  - ***learn* from training examples and exhibits some capability for generalization *beyond* the training examples.**
- **Parallelism and Connectivity**
  - **Great potential for *parallelism* units computations are largely *independent* of each other**
  - ***Massive parallelism* and *high connectivity* are** not **necessarily**

Basic concepts (3)

- Machine designed to model the way in which brain performs tasks
  - implemented by electronic devices and/or software (simulation)
  - Learning is the major emphasis of NN
- Massively parallel distributed processor
  - massive interconnection of simple processing units
  - simple processing units store experience and make it available to use
  - knowledge is acquired from environment thru learning process
- Learning Machine
  - modify synaptic weights to obtain design objective
  - modify own topology - neurons die and new one can grow
- Connectionist network - connectionism

Neural Network Definitions(1)

- According to Haykin, *Neural Networks: A Comprehensive Foundation*, *1999:*
  - – A neural network is a massively parallel distributed processor that has a natural propensity for storing experimental knowledge and making it available for use. It resembles the brain in two respects:
    - Knowledge is acquired by the network through a Learning process.
    - Interneuron connection strengths known as synaptic weights are used to store the knowledge.
  - Another definition:
    - Neural networks are parameterized computational nonlinear algorithms for (numerical) data/signal/image processing.
    - Implemented either on a general-purpose computer or into a dedicated hardware.

Neural Network Definitions (2)

- **According to the DARPA Neural Network Study (1988, AFCEA International Press, p. 60):**
  - **A neural network is a system composed of many simple processing elements operating in parallel whose function is determined by network structure, connection strengths, and the processing performed at computing elements or nodes**
- **According to Nigrin (1993), p. 11:**
  - **A neural network is a circuit composed of a very large number of simple processing elements that are neural based. Each element operates only on local information. Furthermore each element operates asynchronously; thus there is no overall system clock.**
- **According to Zurada (1992), p. xv:**
  - **Artificial neural systems, or neural networks, are physical cellular systems which can acquire, store, and utilize experiential knowledge.**

Basic characteristic of biological neurons(1)

- **About six order of magnitude *slower* than silicon logic gates:**
  - **– neurons operates in millisecond range (10^{-3} sec);**
  - **– silicon gates operates in nanosecond range (10^{-9} sec sec).**
- **A function of a biological neuron seems to be much more *complex* than that of a logical gate.**
- **Reasons for brain's slow rate of operation**
  - **A huge number of neurons (10^{11}) and interconnections (60×10^{12}).**
  - **Complex operations done by neurons.**

Basic characteristic of biological neurons(2)

- **Brain is an information processing system:**
  - **highly complex,**
  - **non-linear,**
  - **Parallel.**
- **Brain performs tasks many times faster than fastest digital computers:**
  - **pattern recognition,**
    - **Example:**
      - **Brain: a complex task of perceptual recognition in 200-300 ms;**
      - **Computer: much lesser complexity task can take hours.**
  - **perception,**
  - **motor control.**

Benefits of Neural Networks(1)

- **Power comes from massively parallel distributed structure and learn to generalize**
  - **generalization : ability to produce reasonable output for inputs not encountered during training**
- **NN cannot provide solution by working individually**
  - **Complex problem is decomposed into simple tasks, and each task is assigned to a NN**
  - **Long way to go to build a computer that mimic human brain**

1. Non-linearity

- **interconnection of non-linear neurons is itself non-linear**
- **desirable property if underlying physical mechanism is non-linear**

Benefits of Neural Networks(2)

2. Input-Output Mapping

- input-output mapping is built by learning from examples
  - reduce differences of desired response and actual response
- non-parametric statistical inference
  - estimate arbitrary decision boundaries in input signal space

3. Adaptivity

- adapt synaptic weight to changes of environment
- NN is retrained to deal with minor change in the operating environment
  - change synaptic weights in real-time
- more robust, reliable behavior in non-stationary environment
- Adaptive pattern recognition, Adaptive signal processing, Adaptive control
- stability-plasticity dilemma

Benefits of Neural Networks(3)

4. Evidential Response

- not only selected class label but also confidence
- confidences can be used to reject
  - recognition accuracy vs. reliability (do only you can do)

5. Contextual Information processing

- (contextual) knowledge is presented in the structure
- every neuron is affected by others

6. Fault Tolerance

- performance degrades gracefully under adverse condition
  - catastrophic failure of d-computer

7. VLSI implementability

- massively parallel nature makes it well suited for VLSI implementation

Benefits of Neural Networks(4)

8. Uniformity of Analysis and Design

- **Neuron is common to all NN**
- **share theories and learning algorithms**
- **modular networks can be built thru seamless integration**

9. Neurobiological Analogy

- **living proof of fault tolerant, fast, powerful processing**
- **Neuroscientists see it as a research tool for neurobiological phenomena**
- **Engineers look to neuroscience for new ideas**

Who is concerned with NNs?

- Computer scientists
  - **non- symbolic information processing with neural nets and learning systems in general.**
- Statisticians
  - **flexible, nonlinear regression and classification models.**
- Engineers in many areas, such as
  - **signal processing and automatic control.**
- Cognitive scientists
  - **possible apparatus to describe models of thinking and consciousness (High- level brain function).**
- Neuro-physiologists
  - **describe and explore medium- level brain function (e. g. memory, sensory system, motor system).**
- Physicists
  - **to model phenomena in statistical mechanics and for a lot of other tasks.**

Some Current Artificial Neural Network Applications

- **Brain modeling**
  - **Models of human development – help children with developmental problems**
  - **Simulations of adult performance – aid our understanding of how the brain works**
  - **Neuropsychological models – suggest remedial actions for brain damaged patients**
- **Real world applications**
  - **Financial modeling – predicting stocks, shares, currency exchange rates**
  - **Other time series prediction – climate, weather, airline marketing tactician**
  - **Computer games – intelligent agents, backgammon, first person shooters**
  - **Control systems – autonomous adaptable robots, microwave controllers**
  - **Pattern recognition – speech recognition, hand-writing recognition, sonar signals**
  - **Data analysis – data compression, data mining, PCA, ICA**
  - **Noise reduction – function approximation, ECG noise reduction**
  - **Bioinformatics – protein secondary structure, DNA sequencing**