Introduction to Neural Networks

2nd Year UG, MSc in Computer Science

http://www.cs.bham.ac.uk/~jxb/inn.html

Lecturer: Dr. John A. Bullinaria

http://www.cs.bham.ac.uk/~jxb

Module Administration and Organisation

Introduction to Neural Networks : Lecture 1 (part 1)

1. Aims and Learning Outcomes
2. Assessment
3. Lecture Plan
4. Recommended Books
5. Comments on Mathematical Requirements

Aims and Learning Outcomes

Aims

1. Introduce the main fundamental principles and techniques of neural network systems.
2. Investigate the principal neural network models and applications.

Learning Outcomes

1. Describe the relation between real brains and simple artificial neural network models.
2. Explain and contrast the most common architectures and learning algorithms for Multi-Layer Perceptrons, Radial-Basis Function Networks, Committee Machines, and Kohonen Self-Organising Maps.
3. Discuss the main factors involved in achieving good learning and generalization performance in neural network systems.
4. Identify the main implementational issues for common neural network systems.
5. Evaluate the practical considerations in applying neural networks to real classification and regression problems.

Assessment

From the Introduction to Neural Networks Module Description

70% 2 hour closed book examination

30% continuous assessment by mini-project report

(Resit by written examination only with the continuous assessment mark carried forward.)

Important Dates

There will be lab sessions on 8/9th and 22/23rd November

The continuous assessment deadline will be 12th January

The 2 hour examination will be in May/June

There will be a resit examination in August/September

Lecture Plan

Week	Session 1 Wednesdays 12:00-13:00	Session 2 Thursdays 10:00-11:00
1	Introduction to Neural Networks and their History.	Biological Neurons and Neural Networks. Artificial Neurons.
2	Networks of Artificial Neurons. Single Layer Perceptrons.	Learning and Generalization in Single Layer Perceptrons.
3	Hebbian Learning. Gradient Descent Learning.	The Generalized Delta Rule. Practical Considerations.
4	Learning in Multi-Layer Perceptrons. Back-Propagation.	Learning with Momentum. Conjugate Gradient Learning.
5	Bias and Variance. Under-Fitting and Over-Fitting.	Improving Generalization.
6	Applications of Multi-Layer Perceptrons.	Exercise Session 1
7	Radial Basis Function Networks: Introduction.	Radial Basis Function Networks: Algorithms.
8	Radial Basis Function Networks: Applications.	Committee Machines.
9	Exercise Session 2	Self Organizing Maps: Fundamentals.
10	Self Organizing Maps: Algorithms and Applications.	Learning Vector Quantization (LVQ).
11	Overview of More Advanced Topics.	Exercise Session 3

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.
Fundamentals of Neural Networks	Laurene Fausett	Prentice Hall, 1994	Good intermediate 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.

Comments on Mathematical Requirements

1. The easiest way to formulate and understand neural networks is in terms of mathematical concepts and equations.

2. Once you have the equations, it is fairly straightforward to convert them into C/C++/Java/Fortran/Pascal/MATLAB programs.

3. 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.

4. Don’t be surprised if it takes a while to become familiar with all the new notation – this is normal!

5. You will not be required to perform any mathematical derivations in the examination.

Introduction to Neural Networks and Their History

Introduction to Neural Networks : Lecture 1 (part 2)

1. What are Neural Networks?
2. Why are Artificial Neural Networks Worth Studying?
3. Learning in Neural Networks
4. A Brief History of the Field
5. Artificial Neural Networks compared with Classical Symbolic AI
6. Some Current Artificial Neural Network Applications

What are Neural Networks ?

1. ***Neural Networks*** (NNs) are networks of neurons, for example, as found in real (i.e. biological) brains.

2. ***Artificial Neurons*** are crude approximations of the neurons found in brains. They may be physical devices, or purely mathematical constructs.

3. ***Artificial Neural Networks*** (ANNs) are networks of Artificial Neurons, and hence constitute crude approximations to parts of real brains. They may be physical devices, or simulated on conventional computers.

4. From a practical point of view, an ANN is just a parallel computational system consisting of many simple processing elements connected together in a specific way in order to perform a particular task.

5. One should never lose sight of how crude the approximations are, and how over-simplified our ANNs are compared to real brains.

Why are Artificial Neural Networks worth studying?

1. They are extremely powerful computational devices (Turing equivalent, universal computers).
2. Massive parallelism makes them very efficient.
3. They can learn and generalize from training data – so there is no need for enormous feats of programming.
4. They are particularly fault tolerant – this is equivalent to the "graceful degradation" found in biological systems.
5. They are very noise tolerant – so they can cope with situations where normal symbolic systems would have difficulty.
6. In principle, they can do anything a symbolic/logic system can do, and more. (In practice, getting them to do it can be rather difficult…)

What are Artificial Neural Networks used for?

As with the field of AI in general, there are two basic goals for neural network research:

Brain modelling : The scientific goal of building models of how real brains work. This can potentially help us understand the nature of human intelligence, formulate better teaching strategies, or better remedial actions for brain damaged patients.

Artificial System Building : The engineering goal of building efficient systems for real world applications. This may make machines more powerful, relieve humans of tedious tasks, and may even improve upon human performance.

These should not be thought of as competing goals. We often use exactly the same networks and techniques for both. Frequently progress is made when the two approaches are allowed to feed into each other. There are fundamental differences though, e.g. the need for biological plausibility in brain modelling, and the need for computational efficiency in artificial system building.

Learning in Neural Networks

There are many forms of neural networks. Most operate by passing neural 'activations' through a network of connected neurons.

One of the most powerful features of neural networks is their ability to ***learn*** and ***generalize*** from a set of training data. They adapt the strengths/weights of the connections between neurons so that the final output activations are correct.

There are three broad types of learning:

1. Supervised Learning (i.e. learning with a teacher)
2. Reinforcement learning (i.e. learning with limited feedback)
3. Unsupervised learning (i.e. learning with no help)

This module will study in some detail the most common learning algorithms for the most common types of neural network.

A Brief History of the Field

1943 McCulloch and Pitts proposed the McCulloch-Pitts neuron model

1949 Hebb published his book *The Organization of Behavior*, in which the Hebbian learning rule was proposed.

1958 Rosenblatt introduced the simple single layer networks now called Perceptrons.

1969 Minsky and Papert's book *Perceptrons* demonstrated the limitation of single layer perceptrons, and almost the whole field went into hibernation.

1982 Hopfield published a series of papers on Hopfield networks.

1982 Kohonen developed the Self-Organising Maps that now bear his name.

1986 The Back-Propagation learning algorithm for Multi-Layer Perceptrons was re-discovered and the whole field took off again.

1990s The sub-field of Radial Basis Function Networks was developed.

2000s The power of Ensembles of Neural Networks and Support Vector Machines becomes apparent.

ANNs compared with Classical Symbolic AI

The distinctions can put under three headings:

1. Level of Explanation
2. Processing Style
3. Representational Structure

These lead to a traditional set of dichotomies:

1. Sub-symbolic vs. Symbolic
2. Non-modular vs. Modular
3. Distributed representation vs. Localist representation
4. Bottom up vs. Top Down
5. Parallel processing vs. Sequential processing

In practice, the distinctions are becoming increasingly blurred.

Some Current Artificial Neural Network Applications

Brain modelling

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 modelling – 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, GTM

Noise reduction – function approximation, ECG noise reduction

Bioinformatics – protein secondary structure, DNA sequencing

Overview and Reading

1. Artificial Neural Networks are powerful computational systems consisting of many simple processing elements connected together to perform tasks analogously to biological brains.
2. They are massively parallel, which makes them efficient, robust, fault tolerant and noise tolerant.
3. They can learn from training data and generalize to new situations.
4. They are useful for brain modelling and real world applications involving pattern recognition, function approximation, prediction, …

Reading

1. Haykin: Sections 1.1, 1.8, 1.9
2. Gurney: Sections 1.1, 1.2, 1.3
3. Beale & Jackson: Sections 1.1, 1.2, 1.3, 1.4
4. Ham & Kostanic: Sections 1.1, 1.2