```
Archive-name: ai-faq/neural-nets/part4
Last-modified: 2002-08-30
URL: ftp://ftp.sas.com/pub/neural/FAQ4.html
Maintainer: saswss@unx.sas.com (Warren S. Sarle)
```

Copyright 1997, 1998, 1999, 2000, 2001, 2002 by Warren S. Sarle, Cary, NC, USA. Reviews provided by other authors as cited below are copyrighted by those authors, who by submitting the reviews for the FAQ give permission for the review to be reproduced as part of the FAQ in any of the ways specified in part 1 of the FAQ.

This is part 4 (of 7) of a monthly posting to the Usenet newsgroup comp.ai.neural-nets. See the part 1 of this posting for full information what it is all about.

========== Questions ==========

Subject: Books and articles about Neural Networks?

The following search engines will search many bookstores for new and used books and return information on availability, price, and shipping charges:

> AddAll: http://www.addall.com/
> Bookfinder: http://www.bookfinder.com/

Clicking on the author and title of most of the books listed in the "Best" and "Notable" sections will do a search using AddAll.

There are many on-line bookstores, such as:

> Amazon: http://www.amazon.com/
> Amazon, UK: http://www.amazon.co.uk/
> Amazon, Germany: http://www.amazon.de/
> Barnes & Noble: http://www.bn.com/
> Bookpool: http://www.bookpool.com/
> Borders: http://www.borders.com/
> Fatbrain: http://www.fatbrain.com/

The neural networks reading group at the University of Illinois at Urbana-Champaign, the Artifical Neural Networks and Computational Brain Theory (ANNCBT) forum, has compiled a large number of book and paper reviews at http://anncbt.ai.uiuc.edu/, with an emphasis more on cognitive science rather than practical applications of NNs.

The Best

The best of the best

Bishop (1995) is clearly the single best book on artificial NNs. This book excels in organization and choice of material, and is a close runner-up to Ripley (1996) for accuracy. If you are new to the field, read it from cover to cover. If you have lots of experience with NNs, it's an excellent reference. If you don't know calculus, take a class. I hope a second edition comes out soon! For more information, see The best intermediate textbooks on NNs below.

If you have questions on feedforward nets that aren't answered by Bishop, try Masters (1993) or Reed and Marks (1999) for practical issues or Ripley (1996) for theortical issues, all of which are reviewed below.

The best popular introduction to NNs

Hinton, G.E. (1992), "How Neural Networks Learn from Experience", Scientific American, 267 (September), 144-151.
Author's Webpage: http://www.cs.utoronto.ca/DCS/People/Faculty/hinton.html (official)
and http://www.cs.toronto.edu/~hinton (private)
Journal Webpage: http://www.sciam.com/

Additional Information: Unfortunately that article is not available there.

The best introductory book for business executives

Bigus, J.P. (1996), *Data Mining with Neural Networks: Solving Business Problems--from Application Development to Decision Support,* NY: McGraw-Hill, ISBN 0-07-005779-6, xvii+221 pages.
The stereotypical business executive (SBE) does not want to know how or why NNs work--he (SBEs are usually male) just wants to make money. The SBE may know what an average or percentage is, but he is deathly afraid of "statistics". He understands profit and loss but does not want to waste his time learning things involving complicated math, such as high-school algebra. For further information on the SBE, see the "Dilbert" comic strip.

Bigus has written an excellent introduction to NNs for the SBE. Bigus says (p. xv), "For business executives, managers, or computer professionals, this book provides a thorough introduction to neural network technology and the issues related to its application without getting bogged down in complex math or needless details. The reader will be able to identify common business problems that are amenable to the neural netwrk approach and will be sensitized to the issues that can affect successful completion of such applications." Bigus succeeds in explaining NNs at a practical, intuitive, and necessarily shallow level without formulas--just what the SBE needs. This book is far better than Caudill and Butler (1990), a popular but disastrous attempt to explain NNs without formulas.

Chapter 1 introduces data mining and data warehousing, and sketches some applications thereof. Chapter 2 is the semi-obligatory philosophico-historical discussion of AI and NNs and is well-written, although the SBE in a hurry may want to skip it. Chapter 3 is a very useful discussion of data preparation. Chapter 4 describes a variety of NNs and what they are good for. Chapter 5 goes into practical issues of training and testing NNs. Chapters 6 and 7 explain how to use the results from NNs. Chapter 8 discusses intelligent agents. Chapters 9 through 12 contain case histories of NN applications, including market segmentation, real-estate pricing, customer ranking, and sales forecasting.

Bigus provides generally sound advice. He briefly discusses overfitting and overtraining without going into much detail, although I think his advice on p. 57 to have at least two training cases for each connection is somewhat lenient, even for noise-free data. I do not understand his claim on pp. 73 and 170 that RBF networks have advantages over backprop networks for nonstationary inputs-- perhaps he is using the word "nonstationary" in a sense different from the statistical meaning of the term. There are other things in the book that I would quibble with, but I did not find any of the flagrant errors that are common in other books on NN applications such as Swingler (1996).

The one serious drawback of this book is that it is more than one page long and may therefore tax the attention span of the SBE. But any SBE who succeeds in reading the entire book should learn enough to be able to hire a good NN expert to do the real work.

The best elementary textbooks

Fausett, L. (1994), *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications,* Englewood Cliffs, NJ: Prentice Hall, ISBN 0-13-334186-0. Also published as a Prentice Hall International Edition, ISBN 0-13-042250-9. Sample software (source code listings in C and Fortran) is included in an Instructor's Manual.
Book Webpage (Publisher): http://www.prenhall.com/books/esm_0133341860.html
Additional Information: The mentioned programs / additional support is not available. Contents:
Ch. 1 Introduction, 1.1 Why Neural Networks and Why Now?, 1.2 What Is a Neural Net?, 1.3 Where Are Neural Nets Being Used?, 1.4 How Are Neural Networks Used?, 1.5 Who Is Developing Neural Networks?, 1.6 When Neural Nets Began: the McCulloch-Pitts Neuron;
Ch. 2 Simple Neural Nets for Pattern Classification, 2.1 General Discussion, 2.2 Hebb Net, 2.3 Perceptron, 2.4 Adaline;
Ch. 3 Pattern Association, 3.1 Training Algorithms for Pattern Association, 3.2 Heteroassociative Memory Neural Network, 3.3 Autoassociative Net, 3.4 Iterative Autoassociative Net, 3.5 Bidirectional Associative Memory (BAM);
Ch. 4 Neural Networks Based on Competition, 4.1 Fixed-Weight Competitive Nets, 4.2 Kohonen Self-Organizing Maps, 4.3 Learning Vector Quantization, 4.4 Counterpropagation;
Ch. 5 Adaptive Resonance Theory, 5.1 Introduction, 5.2 Art1, 5.3 Art2;
Ch. 6 Backpropagation Neural Net, 6.1 Standard Backpropagation, 6.2 Variations, 6.3 Theoretical Results;
Ch. 7 A Sampler of Other Neural Nets, 7.1 Fixed Weight Nets for Constrained Optimization, 7.2 A Few More Nets that Learn, 7.3 Adaptive Architectures, 7.4 Neocognitron; Glossary.

Review by Ian Cresswell:

> What a relief! As a broad introductory text this is without any doubt the best currently available in its area. It doesn't include source code of any kind (normally this is badly written and compiler specific). The algorithms for many different kinds of simple neural nets are presented in a clear step by step manner in plain English.

> Equally, the mathematics is introduced in a relatively gentle manner. There are no unnecessary complications or diversions from the main theme.

The examples that are used to demonstrate the various algorithms are detailed but (perhaps necessarily) simple.

There are bad things that can be said about most books. There are only a small number of minor criticisms that can be made about this one. More space should have been given to backprop and its variants because of the practical importance of such methods. And while the author discusses early stopping in one paragraph, the treatment of generalization is skimpy compared to the books by Weiss and Kulikowski or Smith listed above.

If you're new to neural nets and you don't want to be swamped by bogus ideas, huge amounts of intimidating looking mathematics, a programming language that you don't know etc. etc. then this is the book for you.

In summary, this is the best starting point for the outsider and/or beginner... a truly excellent text.

Smith, M. (1996). *Neural Networks for Statistical Modeling,* NY: Van Nostrand Reinhold, ISBN 0-442-01310-8.
Apparently there is a new edition I haven't seen yet:
Smith, M. (1996). *Neural Networks for Statistical Modeling,* Boston: International Thomson Computer Press, ISBN 1-850-32842-0.
Book Webpage (Publisher): http://www.thompson.com/
Publisher's address: 20 Park Plaza, Suite 1001, Boston, MA 02116, USA.
Smith is not a statistician, but he has made an impressive effort to convey statistical fundamentals applied to neural networks. The book has entire brief chapters on overfitting and validation (early stopping and split-sample validation, which he incorrectly calls cross-validation), putting it a rung above most other introductions to NNs. There are also brief chapters on data preparation and diagnostic plots, topics usually ignored in elementary NN books. Only feedforward nets are covered in any detail.
Chapter headings: Mapping Functions; Basic Concepts; Error Derivatives; Learning Laws; Weight Initialization; The Course of Learning: An Example; Overfitting; Cross Validation; Preparing the Data; Representing Variables; Using the Model.

Weiss, S.M. and Kulikowski, C.A. (1991), *Computer Systems That Learn,* Morgan Kaufmann. ISBN 1-55860-065-5.
Author's Webpage: Kulikowski: http://ruccs.rutgers.edu/faculty/kulikowski.html
Book Webpage (Publisher): http://www.mkp.com/books_catalog/1-55860-065-5.asp
Additional Information: Information of Weiss, S.M. are not available.
Briefly covers at a very elementary level feedforward nets, linear and nearest-neighbor discriminant analysis, trees, and expert sytems, emphasizing practical applications. For a book at this level, it has an unusually good chapter on estimating generalization error, including bootstrapping.

```
1 Overview of Learning Systems
    1.1 What is a Learning System?
    1.2 Motivation for Building Learning Systems
    1.3 Types of Practical Empirical Learning Systems
        1.3.1 Common Theme: The Classification Model
        1.3.2 Let the Data Speak
    1.4 What's New in Learning Methods
        1.4.1 The Impact of New Technology
    1.5 Outline of the Book
    1.6 Bibliographical and Historical Remarks

2 How to Estimate the True Performance of a Learning System
    2.1 The Importance of Unbiased Error Rate Estimation
    2.2. What is an Error?
        2.2.1 Costs and Risks
    2.3 Apparent Error Rate Estimates
    2.4 Too Good to Be True: Overspecialization
    2.5 True Error Rate Estimation
        2.5.1 The Idealized Model for Unlimited Samples
        2.5.2 Train-and Test Error Rate Estimation
        2.5.3 Resampling Techniques
        2.5.4 Finding the Right Complexity Fit
    2.6 Getting the Most Out of the Data
    2.7 Classifier Complexity and Feature Dimensionality
        2.7.1 Expected Patterns of Classifier Behavior
    2.8 What Can Go Wrong?
        2.8.1 Poor Features, Data Errors, and Mislabeled Classes
        2.8.2 Unrepresentative Samples
    2.9 How Close to the Truth?
    2.10 Common Mistakes in Performance Analysis
    2.11 Bibliographical and Historical Remarks

3 Statistical Pattern Recognition
    3.1 Introduction and Overview
    3.2 A Few Sample Applications
    3.3 Bayesian Classifiers
        3.3.1 Direct Application of the Bayes Rule
    3.4 Linear Discriminants
        3.4.1 The Normality Assumption and Discriminant Functions
        3.4.2 Logistic Regression
    3.5 Nearest Neighbor Methods
    3.6 Feature Selection
    3.7 Error Rate Analysis
```

Reed, R.D., and Marks, R.J, II (1999), *Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks,*
Cambridge, MA: The MIT Press, ISBN 0-262-18190-8.
Author's Webpage: Marks: http://cialab.ee.washington.edu/Marks.html
Book Webpage (Publisher): http://mitpress.mit.edu/book-home.tcl?isbn=0262181908
After you have read Smith (1996) or Weiss and Kulikowski (1991), consult Reed and Marks for practical details on training MLPs (other types of neural nets such as RBF networks are barely even mentioned). They provide extensive coverage of backprop and its variants, and they also survey conventional optimization algorithms. Their coverage of initialization methods, constructive networks, pruning, and regularization methods is unusually thorough. Unlike the vast majority of books on neural nets, this one has lots of really informative graphs. The chapter on generalization assessment is slightly weak, which is why you should read Smith (1996) or Weiss and Kulikowski (1991) first. Also, there is little information on data preparation, for which Smith (1996) and Masters (1993; see below) should be consulted. There is some elementary calculus, but not enough that it should scare off anybody. Many second-rate books treat neural nets as mysterious black boxes, but Reed and Marks open up the box and provide genuine insight into the way neural nets work.

One problem with the book is that the terms "validation set" and "test set" are used inconsistently.

Chapter headings: Supervised Learning; Single-Layer Networks; MLP Representational Capabilities; Back-Propagation; Learning Rate and Momentum; Weight-Initialization Techniques; The Error Surface; Faster Variations of Back-Propagation; Classical Optimization Techniques; Genetic Algorithms and Neural Networks; Constructive Methods; Pruning Algorithms; Factors Influencing Generalization; Generalization Prediction and Assessment; Heuristics for Improving Generalization; Effects of Training with Noisy Inputs; Linear Regression; Principal Components Analysis; Jitter Calculations; Sigmoid-like Nonlinear Functions

The best books on using and programming NNs

Masters, T. (1993), *Practical Neural Network Recipes in C++,* Academic Press, ISBN 0-12-479040-2, US $45 incl. disks.
Book Webpage (Publisher): http://www.apcatalog.com/cgi-bin/AP?ISBN=0124790402&LOCATION=US&FORM=FORM2
Masters has written three exceptionally good books on NNs (the two others are listed below). He combines generally sound practical advice with some basic statistical knowledge to produce a programming text that is far superior to the competition (see "The Worst"

below). Not everyone likes his C++ code (the usual complaint is that the code is not sufficiently OO) but, unlike the code in some other books, Masters's code has been successfully compiled and run by some readers of comp.ai.neural-nets. Masters's books are well worth reading even for people who have no interest in programming.

Chapter headings: Foundations; Classification; Autoassociation; Time-Series Prediction; Function Approximation; Multilayer Feedforward Networks; Eluding Local Minima I: Simulated Annealing; Eluding Local Minima II: Genetic Optimization; Regression and Neural Networks; Designing Feedforward Network Architectures; Interpreting Weights: How Does This Thing Work; Probabilistic Neural Networks; Functional Link Networks; Hybrid Networks; Designing the Training Set; Preparing Input Data; Fuzzy Data and Processing; Unsupervised Training; Evaluating Performance of Neural Networks; Confidence Measures; Optimizing the Decision Threshold; Using the NEURAL Program.

Masters, T. (1995) *Advanced Algorithms for Neural Networks: A C++ Sourcebook,* NY: John Wiley and Sons, ISBN 0-471-10588-0
Book Webpage (Publisher): http://www.wiley.com/
Additional Information: One has to search.
Clear explanations of conjugate gradient and Levenberg-Marquardt optimization algorithms, simulated annealing, kernel regression (GRNN) and discriminant analysis (PNN), Gram-Charlier networks, dimensionality reduction, cross-validation, and bootstrapping.

Masters, T. (1994), *Signal and Image Processing with Neural Networks: A C++ Sourcebook,* NY: Wiley, ISBN 0-471-04963-8.
Book Webpage (Publisher): http://www.wiley.com/
Additional Information: One has to search.

The best intermediate textbooks on NNs

Bishop, C.M. (1995). *Neural Networks for Pattern Recognition,* Oxford: Oxford University Press. ISBN 0-19-853849-9 (hardback) or 0-19-853864-2 (paperback), xvii+482 pages.
Author's Webpage: http://neural-server.aston.ac.uk/People/bishopc/Welcome.html
Book Webpage (Publisher): http://www.oup.co.uk/isbn/0-19-853864-2
This is definitely the best book on feedforward neural nets for readers comfortable with calculus. The book is exceptionally well organized, presenting topics in a logical progression ideal for conceptual understanding.

Geoffrey Hinton writes in the foreword:
"Bishop is a leading researcher who has a deep understanding of the material and has gone to great lengths to organize it in a sequence that makes sense. He has wisely avoided the temptation to try to cover everything and has therefore omitted interesting topics like reinforcement learning, Hopfield networks, and Boltzmann machines in order to focus on the types of neural networks that are most widely used in practical applications. He assumes that the reader has the basic mathematical literacy required for an undergraduate science degree, and using these tools he explains everything from scratch. Before introducing the multilayer perceptron, for example, he lays a solid foundation of basic statistical concepts. So the crucial concept of overfitting is introduced using easily visualized examples of one-dimensional polynomials and only later applied to neural networks. An impressive aspect of this book is that it takes the reader all the way from the simplest linear models to the very latest Bayesian multilayer neural networks without ever requiring any great intellectual leaps."

Chapter headings: Statistical Pattern Recognition; Probability Density Estimation; Single-Layer Networks; The Multi-layer Perceptron; Radial Basis Functions; Error Functions; Parameter Optimization Algorithms; Pre-processing and Feature Extraction; Learning and Generalization; Bayesian Techniques; Symmetric Matrices; Gaussian Integrals; Lagrange Multipliers; Calculus of Variations; Principal Components.

Hertz, J., Krogh, A., and Palmer, R. (1991). *Introduction to the Theory of Neural Computation.* Redwood City, CA: Addison-Wesley, ISBN 0-201-50395-6 (hardbound) and 0-201-51560-1 (paperbound)
Book Webpage (Publisher): http://www2.awl.com/gb/abp/sfi/computer.html
This is an excellent classic work on neural nets from the perspective of physics covering a wide variety of networks. Comments from readers of comp.ai.neural-nets: "My first impression is that this one is by far the best book on the topic. And it's below $30 for the paperback."; "Well written, theoretical (but not overwhelming)"; It provides a good balance of model development, computational algorithms, and applications. The mathematical derivations are especially well done"; "Nice mathematical analysis on the mechanism of different learning algorithms"; "It is NOT for mathematical beginner. If you don't have a good grasp of higher level math, this book can be really tough to get through."

The best advanced textbook covering NNs

Ripley, B.D. (1996) *Pattern Recognition and Neural Networks,* Cambridge: Cambridge University Press, ISBN 0-521-46086-7 (hardback), xii+403 pages.
Author's Webpage: http://www.stats.ox.ac.uk/~ripley/
Book Webpage (Publisher): http://www.cup.cam.ac.uk/
Additional Information: The Webpage includes errata and additional information, which hasn't been available at publishing time, for this book.
Brian Ripley's book is an excellent sequel to Bishop (1995). Ripley starts up where Bishop left off, with Bayesian inference and

statistical decision theory, and then covers some of the same material on NNs as Bishop but at a higher mathematical level. Ripley also covers a variety of methods that are not discussed, or discussed only briefly, by Bishop, such as tree-based methods and belief networks. While Ripley is best appreciated by people with a background in mathematical statistics, the numerous realistic examples in his book will be of interest even to beginners in neural nets.

Chapter headings: Introduction and Examples; Statistical Decision Theory; Linear Discriminant Analysis; Flexible Discriminants; Feed-forward Neural Networks; Non-parametric Methods; Tree-structured Classifiers; Belief Networks; Unsupervised Methods; Finding Good Pattern Features; Statistical Sidelines.

Devroye, L., Györfi, L., and Lugosi, G. (1996), *A Probabilistic Theory of Pattern Recognition,* NY: Springer, ISBN 0-387-94618-7, vii+636 pages.

This book has relatively little material explicitly about neural nets, but what it has is very interesting and much of it is not found in other texts. The emphasis is on statistical proofs of universal consistency for a wide variety of methods, including histograms, (k) nearest neighbors, kernels (PNN), trees, generalized linear discriminants, MLPs, and RBF networks. There is also considerable material on validation and cross-validation. The authors say, "We did not scar the pages with backbreaking simulations or quick-and-dirty engineering solutions" (p. 7). The formula-to-text ratio is high, but the writing is quite clear, and anyone who has had a year or two of mathematical statistics should be able to follow the exposition.

Chapter headings: The Bayes Error; Inequalities and Alternate Distance Measures; Linear Discrimination; Nearest Neighbor Rules; Consistency; Slow Rates of Convergence; Error Estimation; The Regular Histogram Rule; Kernel Rules; Consistency of the k-Nearest Neighbor Rule; Vapnik-Chervonenkis Theory; Combinatorial Aspects of Vapnik-Chervonenkis Theory; Lower Bounds for Empirical Classifier Selection; The Maximum Likelihood Principle; Parametric Classification; Generalized Linear Discrimination; Complexity Regularization; Condensed and Edited Nearest Neighbor Rules; Tree Classifiers; Data-Dependent Partitioning; Splitting the Data; The Resubstitution Estimate; Deleted Estimates of the Error Probability; Automatic Kernel Rules; Automatic Nearest Neighbor Rules; Hypercubes and Discrete Spaces; Epsilon Entropy and Totally Bounded Sets; Uniform Laws of Large Numbers; Neural Networks; Other Error Estimates; Feature Extraction.

The best books on neurofuzzy systems

Brown, M., and Harris, C. (1994), *Neurofuzzy Adaptive Modelling and Control,* NY: Prentice Hall, ISBN 0-13-134453-6.
Author's Webpage: http://www.isis.ecs.soton.ac.uk/people/m_brown.html
and http://www.ecs.soton.ac.uk/~cjh/
Book Webpage (Publisher): http://www.prenhall.com/books/esm_0131344536.html
Additional Information: Additional page at: http://www.isis.ecs.soton.ac.uk/publications/neural/mqbcjh94e.html and an abstract can be found at: http://www.isis.ecs.soton.ac.uk/publications/neural/mqb93.html
Brown and Harris rely on the fundamental insight that that a fuzzy system is a nonlinear mapping from an input space to an output space that can be parameterized in various ways and therefore can be adapted to data using the usual neural training methods (see "What is backprop?") or conventional numerical optimization algorithms (see "What are conjugate gradients, Levenberg-Marquardt, etc.?"). Their approach makes clear the intimate connections between fuzzy systems, neural networks, and statistical methods such as B-spline regression.

The best comparison of NNs with other classification methods

Michie, D., Spiegelhalter, D.J. and Taylor, C.C. (1994), *Machine Learning, Neural and Statistical Classification*, Ellis Horwood.
Author's Webpage: Donald Michie: http://www.aiai.ed.ac.uk/~dm/dm.html
Additional Information: This book is out of print but available online at http://www.amsta.leeds.ac.uk/~charles/statlog/

Other notable books

Introductory

Anderson, J.A. (1995), *An Introduction to Neural Networks,* Cambridge,MA: The MIT Press, ISBN 0-262-01144-1.
Author's Webpage: http://www.cog.brown.edu/~anderson
Book Webpage (Publisher): http://mitpress.mit.edu/book-home.tcl?isbn=0262510812 or
http://mitpress.mit.edu/book-home.tcl?isbn=0262011441 (hardback)
Additional Information: Programs and additional information can be found at: ftp://mitpress.mit.edu/pub/Intro-to-NeuralNets/
Anderson provides an accessible introduction to the AI and neurophysiological sides of NN research, although the book is weak regarding practical aspects of using NNs.
Chapter headings: Properties of Single Neurons; Synaptic Integration and Neuron Models; Essential Vector Operations; Lateral Inhibition and Sensory Processing; Simple Matrix Operations; The Linear Associator: Background and Foundations; The Linear Associator: Simulations; Early Network Models: The Perceptron; Gradient Descent Algorithms; Representation of Information; Applications of Simple Associators: Concept Formation and Object Motion; Energy and Neural Networks: Hopfield Networks and Boltzmann Machines; Nearest Neighbor Models; Adaptive Maps; The BSB Model: A Simple Nonlinear Autoassociative Neural Network; Associative Computation; Teaching Arithmetic to a Neural Network.

Hagan, M.T., Demuth, H.B., and Beale, M. (1996), *Neural Network Design,* Boston: PWS, ISBN 0-534-94332-2.

It doesn't really say much about design, but this book provides formulas and examples in excruciating detail for a wide variety of networks. It also includes some mathematical background material.
Chapter headings: Neuron Model and Network Architectures; An Illustrative Example; Perceptron Learning Rule; Signal and Weight Vector Spaces; Linear Transformations for Neural; Networks; Supervised Hebbian Learning; Performance Surfaces and Optimum Points; Performance Optimization; Widrow-Hoff Learning; Backpropagation; Variations on Backpropagation; Associative Learning; Competitive Networks; Grossberg Network; Adaptive Resonance Theory; Stability; Hopfield Network.

Abdi, H., Valentin, D., and Edelman, B. (1999), *Neural Networks,* Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-124, Thousand Oaks, CA: Sage, ISBN 0-7619-1440-4.
Inexpensive, brief (89 pages) but very detailed explanations of linear networks and the basics of backpropagation.
Chapter headings: 1. Introduction 2. The Perceptron 3. Linear Autoassociative Memories 4. Linear Heteroassociative Memories 5. Error Backpropagation 6. Useful References.

Bayesian learning

Neal, R. M. (1996) *Bayesian Learning for Neural Networks,* New York: Springer-Verlag, ISBN 0-387-94724-8.

Biological learning and neurophysiology

Koch, C., and Segev, I., eds. (1998) *Methods in Neuronal Modeling: From Ions to Networks,* 2nd ed., Cambridge, MA: The MIT Press, ISBN 0-262-11231-0.
Book Webpage: http://goethe.klab.caltech.edu/MNM/

Rolls, E.T., and Treves, A. (1997), *Neural Networks and Brain Function,* Oxford: Oxford University Press, ISBN: 0198524323.
Chapter headings: Introduction; Pattern association memory; Autoassociation memory; Competitive networks, including self-organizing maps; Error-correcting networks: perceptrons, the delta rule, backpropagation of error in multilayer networks, and reinforcement learning algorithms; The hippocampus and memory; Pattern association in the brain: amygdala and orbitofrontal cortex; Cortical networks for invariant pattern recognition; Motor systems: cerebellum and basal ganglia; Cerebral neocortex.

Schmajuk, N.A. (1996) *Animal Learning and Cognition: A Neural Network Approach,* Cambridge: Cambridge University Press, ISBN 0521456967.
Chapter headings: Neural networks and associative learning Classical conditioning: data and theories; Cognitive mapping; Attentional processes; Storage and retrieval processes; Configural processes; Timing; Operant conditioning and animal communication: data, theories, and networks; Animal cognition: data and theories; Place learning and spatial navigation; Maze learning and cognitive mapping; Learning, cognition, and the hippocampus: data and theories; Hippocampal modulation of learning and cognition; The character of the psychological law.

Collections

Orr, G.B., and Mueller, K.-R., eds. (1998), *Neural Networks: Tricks of the Trade,* Berlin: Springer, ISBN 3-540-65311-2.
Articles: Efficient BackProp; Early Stopping - But When? A Simple Trick for Estimating the Weight Decay Parameter; Controling the Hyperparameter Search in MacKay's Bayesian Neural Network Framework; Adaptive Regularization in Neural Network Modeling; Large Ensemble Averaging; Square Unit Augmented, Radially Extended, Multilayer Perceptrons; A Dozen Tricks with Multitask Learning; Solving the Ill-Conditioning in Neural Network Learning; Centering Neural Network Gradient Factors; Avoiding Roundoff Error in Backpropagating Derivatives; Transformation Invariance in Pattern Recognition - Tangent Distance and Tangent Propagation; Combining Neural Networks and Context-Driven Search for On-Line, Printed Handwriting Recognition in the Newton; Neural Network Classification and Prior Class Probabilities; Applying Divide and Conquer to Large Scale Pattern Recognition Tasks; Forecasting the Economy with Neural Nets: A Survey of Challenges and Solutions; How to Train Neural Networks.

Arbib, M.A., ed. (1995), *The Handbook of Brain Theory and Neural Networks,* Cambridge, MA: The MIT Press, ISBN 0-262-51102-9.
From The Publisher: The heart of the book, part III, comprises of 267 original articles by leaders in the various fields, arranged alphabetically by title. Parts I and II, written by the editor, are designed to help readers orient themselves to this vast range of material. Part I, Background, introduces several basic neural models, explains how the present study of brain theory and neural networks integrates brain theory, artificial intelligence, and cognitive psychology, and provides a tutorial on the concepts essential for understanding neural networks as dynamic, adaptive systems. Part II, Road Maps, provides entry into the many articles of part III through an introductory "Meta-Map" and twenty-three road maps, each of which tours all the Part III articles on the chosen theme.

Touretzky, D., Hinton, G, and Sejnowski, T., eds., (1989) *Proceedings of the 1988 Connectionist Models Summer School,* San Mateo, CA: Morgan Kaufmann, ISBN: 1558600337

NIPS:

1. Touretzky, D.S., ed. (1989), *Advances in Neural Information Processing Systems 1,* San Mateo, CA: Morgan Kaufmann, ISBN:

1558600159

2. Touretzky, D. S., ed. (1990), *Advances in Neural Information Processing Systems 2,* San Mateo, CA: Morgan Kaufmann, ISBN: 1558601007
3. Lippmann, R.P., Moody, J.E., and Touretzky, D. S., eds. (1991) *Advances in Neural Information Processing Systems 3,* San Mateo, CA: Morgan Kaufmann, ISBN: 1558601848
4. Moody, J.E., Hanson, S.J., and Lippmann, R.P., eds. (1992) *Advances in Neural Information Processing Systems 4,* San Mateo, CA: Morgan Kaufmann, ISBN: 1558602224
5. Hanson, S.J., Cowan, J.D., and Giles, C.L. eds. (1993) *Advances in Neural Information Processing Systems 5,* San Mateo, CA: Morgan Kaufmann, ISBN: 1558602747
6. Cowan, J.D., Tesauro, G., and Alspector, J., eds. (1994) *Advances in Neural Information Processing Systems 6,* San Mateo, CA: Morgan Kaufman, ISBN: 1558603220
7. Tesauro, G., Touretzky, D., and Leen, T., eds. (1995) *Advances in Neural Information Processing Systems 7,* Cambridge, MA: The MIT Press, ISBN: 0262201046
8. Touretzky, D. S., Mozer, M.C., and Hasselmo, M.E., eds. (1996) *Advances in Neural Information Processing Systems 8,* Cambridge, MA: The MIT Press, ISBN: 0262201070
9. Mozer, M.C., Jordan, M.I., and Petsche, T., eds. (1997) *Advances in Neural Information Processing Systems 9,* Cambridge, MA: The MIT Press, ISBN: 0262100657
10. Jordan, M.I., Kearns, M.S., and Solla, S.A., eds. (1998) *Advances in Neural Information Processing Systems 10,* Cambridge, MA: The MIT Press, ISBN: 0262100762
11. Kearns, M.S., Solla, S.A., amd Cohn, D.A., eds. (1999) *Advances in Neural Information Processing Systems 11,* Cambridge,MA: The MIT Press, ISBN: 0262112450
12. Solla, S.A., Leen, T., and Müller, K.-R., eds. (2000) *Advances in Neural Information Processing Systems 12,* Cambridge, MA: The MIT Press, ISBN: 0-262-19450-3

Combining networks

Sharkey, A.J.C. (1999), *Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems,* London: Springer, ISBN: 185233004X

Connectionism

Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., and Parisi, D. (1996) *Rethinking Innateness: A Connectionist Perspective on Development,* Cambridge, MA: The MIT Press, ISBN: 026255030X.
Chapter headings: New perspectives on development; Why connectionism? Ontogenetic development: A connectionist synthesis; The shape of change; Brain development; Interactions, all the way down; Rethinking innateness.

Plunkett, K., and Elman, J.L. (1997), *Exercises in Rethinking Innateness: A Handbook for Connectionist Simulations,* Cambridge, MA: The MIT Press, ISBN: 0262661055.
Chapter headings: Introduction and overview; The methodology of simulations; Learning to use the simulator; Learning internal representations; Autoassociation; Generalization; Translation invariance; Simple recurrent networks; Critical points in learning; Modeling stages in cognitive development; Learning the English past tense; The importance of starting small.

Feedforward networks

Fine, T.L. (1999) *Feedforward Neural Network Methodology,* NY: Springer, ISBN 0-387-98745-2.

Husmeier, D. (1999), *Neural Networks for Conditional Probability Estimation: Forecasting Beyond Point Predictions,* Berlin: Springer Verlag, ISBN 185233095.

Fuzzy logic and neurofuzzy systems

See also "General (including SVMs and Fuzzy Logic)".

Kosko, B. (1997), *Fuzzy Engineering,* Upper Saddle River, NJ: Prentice Hall, ISBN 0-13-124991-6.
Kosko's new book is a big improvement over his older neurofuzzy book and makes an excellent sequel to Brown and Harris (1994).

Nauck, D., Klawonn, F., and Kruse, R. (1997), *Foundations of Neuro-Fuzzy Systems,* Chichester: Wiley, ISBN 0-471-97151-0.
Chapter headings: Historical and Biological Aspects; Neural Networks; Fuzzy Systems; Modelling Neuro-Fuzzy Systems; Cooperative Neuro-Fuzzy Systems; Hybrid Neuro-Fuzzy Systems; The Generic Fuzzy Perceptron; NEFCON - Neuro-Fuzzy Control; NEFCLASS - Neuro-Fuzzy Classification; NEFPROX - Neuro-Fuzzy Function Approximation; Neural Networks and Fuzzy Prolog; Using Neuro-Fuzzy Systems.

General (including SVMs and Fuzzy Logic)

Many books on neural networks, machine learning, etc., present various methods as miscellaneous tools without any conceptual framework relating different methods. The best of such neural net "cookbooks" is probably Haykin's (1999) second edition.

Among conceptually-integrated books, there are two excellent books that use the Vapnil-Chervonenkis theory as a unifying theme, and provide strong coverage of support vector machines and fuzzy logic, as well as neural nets. Of these two, Kecman (2001) provides clearer explanations and better diagrams, but Cherkassky and Mulier (1998) are better organized have an excellent section on unsupervised learning, especially self-organizing maps. I have been tempted to add both of these books to the "best" list, but I have not done so because I think VC theory is of doubtful practical utility for neural nets. However, if you are especially interested in VC theory and support vector machines, then both of these books can be highly recommended. To help you choose between them, a detailed table of contents is provided below for each book.

Haykin, S. (1999), *Neural Networks: A Comprehensive Foundation,* 2nd ed., Upper Saddle River, NJ: Prentice Hall, ISBN 0-13-273350-1.
The second edition is much better than the first, which has been described as a core-dump of Haykin's brain. The second edition covers more topics, is easier to understand, and has better examples.
Chapter headings: Introduction; Learning Processes; Single Layer Perceptrons; Multilayer Perceptrons; Radial-Basis Function Networks; Support Vector Machines; Committee Machines; Principal Components Analysis; Self-Organizing Maps; Information-Theoretic Models; Stochastic Machines And Their Approximates Rooted in Statistical Mechanics; Neurodynamic Programming; Temporal Processing Using Feedforward Networks; Neurodynamics; Dynamically Driven Recurrent Networks.

Kecman, V. (2001), *Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models,* Cambridge, MA: The MIT Press; ISBN: 0-262-11255-8.
URL: http://www.support-vector.ws/

```
Detailed Table of Contents:

1. Learning and Soft Computing: Rationale, Motivations, Needs, Basics
    1.1 Examples of Applications in Diverse Fields
    1.2 Basic Tools of Soft Computing: Neural Networks, Fuzzy Logic Systems, and Support Vector Machines
        1.2.1 Basics of Neural Networks
        1.2.2 Basics of Fuzzy Logic Modeling
    1.3 Basic Mathematics of Soft Computing
        1.3.1 Approximation of Multivariate Functions
        1.3.2 Nonlinear Error Surface and Optimization
    1.4 Learning and Statistical Approaches to Regression and Classification
        1.4.1 Regression
        1.4.2 Classification
    Problems
    Simulation Experiments

2. Support Vector Machines
    2.1 Risk Minimization Principles and the Concept of Uniform Convergence
    2.2 The VC Dimension
    2.3 Structural Risk Minimization
    2.4 Support Vector Machine Algorithms
        2.4.1 Linear Maximal Margin Classifier for Linearly Separable Data
        2.4.2 Linear Soft Margin Classifier for Overlapping Classes
        2.4.3 The Nonlinear Classifier
        2.4.4 Regression by Support Vector
    Machines
    Problems
    Simulation Experiments

3. Single-Layer Networks
    3.1 The Perceptron
        3.1.1 The Geometry of Perceptron Mapping
        3.1.2 Convergence Theorem and
    Perceptron Learning Rule
    3.2 The Adaptive Linear Neuron (Adaline) and the Least Mean Square Algorithm
        3.2.1 Representational Capabilities of the Adaline
        3.2.2 Weights Learning for a Linear Processing Unit
    Problems
    Simulation Experiments

4. Multilayer Perceptrons
    4.1  The Error Backpropagation Algorithm
    4.2  The Generalized Delta Rule
    4.3  Heuristics or Practical Aspects of the Error Backpropagation Algorithm
        4.3.1 One, Two, or More Hidden Layers?
        4.3.2 Number of Neurons in a Hidden Layer, or the Bias-Variance Dilemma
        4.3.3 Type of Activation Functions in a Hidden Layer and the Geometry of Approximation
        4.3.4 Weights Initialization
        4.3.5 Error Function for Stopping Criterion at Learning
        4.3.6 Learning Rate and the Momentum Term
    Problems
    Simulation Experiments
```

```
5. Radial Basis Function Networks
    5.1 Ill-Posed Problems and the Regularization Technique
    5.2 Stabilizers and Basis Functions
    5.3 Generalized Radial Basis Function Networks
        5.3.1 Moving Centers Learning
        5.3.2 Regularization with Nonradial Basis Functions
        5.3.3 Orthogonal Least Squares
        5.3.4 Optimal Subset Selection by Linear
    Programming
    Problems
    Simulation Experiments

6. Fuzzy Logic Systems
    6.1 Basics of Fuzzy Logic Theory
        6.1.1 Crisp (or Classic) and Fuzzy Sets
        6.1.2 Basic Set Operations
        6.1.3 Fuzzy Relations
        6.1.4 Composition of Fuzzy Relations
        6.1.5 Fuzzy Inference
        6.1.6 Zadeh's Compositional Rule of Inference
        6.1.7 Defuzzification
    6.2 Mathematical Similarities between Neural Networks and Fuzzy Logic Models
    6.3 Fuzzy Additive Models
    Problems
    Simulation Experiments

7. Case Studies
    7.1 Neural Networks-Based Adaptive Control
        7.1.1 General Learning Architecture, or Direct Inverse Modeling
        7.1.2 Indirect Learning Architecture
        7.1.3 Specialized Learning Architecture
        7.1.4 Adaptive Backthrough Control
    7.2 Financial Time Series Analysis
    7.3 Computer Graphics
        7.3.1 One-Dimensional Morphing
        7.3.2 Multidimensional Morphing
        7.3.3 Radial Basis Function Networks for Human Animation
        7.3.4 Radial Basis Function Networks for Engineering Drawings

8. Basic Nonlinear Optimization Methods
    8.1 Classical Methods
        8.1.1 Newton-Raphson Method
        8.1.2 Variable Metric or Quasi-Newton Methods
        8.1.3 Davidon-Fletcher-Powel Method
        8.1.4 Broyden-Fletcher-Goldfarb-Shano Method
        8.1.5 Conjugate Gradient Methods
        8.1.6 Fletcher-Reeves Method
        8.1.7 Polak-Ribiere Method
        8.1.8 Two Specialized Algorithms for a Sum-of-Error-Squares Error Function
                Gauss-Newton Method
                Levenberg-Marquardt Method
    8.2 Genetic Algorithms and Evolutionary Computing
        8.2.1 Basic Structure of Genetic Algorithms
        8.2.2 Mechanism of Genetic Algorithms

9. Mathematical Tools of Soft Computing
    9.1 Systems of Linear Equations
    9.2 Vectors and Matrices
    9.3 Linear Algebra and Analytic Geometry
    9.4 Basics of Multivariable Analysis
    9.5 Basics from Probability Theory
```

Cherkassky, V.S., and Mulier, F.M. (1998), *Learning from Data : Concepts, Theory, and Methods,* NY: John Wiley & Sons; ISBN: 0-471-15493-8.

```
Detailed Table of Contents:

1 Introduction

   1.1 Learning and Statistical Estimation
   1.2 Statistical Dependency and Causality
   1.3 Characterization of Variables
   1.4 Characterization of Uncertainty
   References

2 Problem Statement, Classical Approaches, and Adaptive Learning

   2.1 Formulation of the Learning Problem
     2.1.1 Role of the Learning Machine
     2.1.2 Common Learning Tasks
     2.1.3 Scope of the Learning Problem Formulation
   2.2 Classical Approaches
     2.2.1 Density Estimation
```

History

Hebb, D.O. (1949), *The Organization of Behavior,* NY: Wiley. Out of print.

Rosenblatt, F. (1962), *Principles of Neurodynamics,* NY: Spartan Books. Out of print.

Anderson, J.A., and Rosenfeld, E., eds. (1988), *Neurocomputing: Foundatons of Research,* Cambridge, MA: The MIT Press, ISBN 0-262-01097-6.
Author's Webpage: http://www.cog.brown.edu/~anderson
Book Webpage (Publisher): http://mitpress.mit.edu/book-home.tcl?isbn=0262510480
43 articles of historical importance, ranging from William James to Rumelhart, Hinton, and Williams.

Anderson, J. A., Pellionisz, A. and Rosenfeld, E. (Eds). (1990). *Neurocomputing 2: Directions for Research.* The MIT Press: Cambridge, MA.
Author's Webpage: http://www.cog.brown.edu/~anderson
Book Webpage (Publisher): http://mitpress.mit.edu/book-home.tcl?isbn=0262510758

Carpenter, G.A., and Grossberg, S., eds. (1991), *Pattern Recognition by Self-Organizing Neural Networks,* Cambridge, MA: The MIT Press, ISBN 0-262-03176-0
Articles on ART, BAM, SOMs, counterpropagation, etc.

Nilsson, N.J. (1965/1990), *Learning Machines,* San Mateo, CA: Morgan Kaufmann, ISBN 1-55860-123-6.

Minsky, M.L., and Papert, S.A. (1969/1988) *Perceptrons,* Cambridge, MA: The MIT Press, 1st ed. 1969, expanded edition 1988 ISBN 0-262-63111-3.

Werbos, P.J. (1994), *The Roots of Backpropagation,* NY: John Wiley & Sons, ISBN: 0471598976. Includes Werbos's 1974 Harvard Ph.D. thesis, *Beyond Regression.*

Kohonen, T. (1984/1989), *Self-organization and Associative Memory,* 1st ed. 1988, 3rd ed. 1989, NY: Springer.
Author's Webpage: http://www.cis.hut.fi/nnrc/teuvo.html
Book Webpage (Publisher): http://www.springer.de/
Additional Information: Book is out of print.

Rumelhart, D. E. and McClelland, J. L. (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition,* Volumes 1 & 2, Cambridge, MA: The MIT Press ISBN 0-262-63112-1.
Author's Webpage: http://www-med.stanford.edu/school/Neurosciences/faculty/rumelhart.html
Book Webpage (Publisher): http://mitpress.mit.edu/book-home.tcl?isbn=0262631121

Hecht-Nielsen, R. (1990), *Neurocomputing,* Reading, MA: Addison-Wesley, ISBN 0-201-09355-3.
Book Webpage (Publisher): http://www.awl.com/

Anderson, J.A., and Rosenfeld, E., eds. (1998), *Talking Nets: An Oral History of Neural Networks,* Cambridge, MA: The MIT Press, ISBN 0-262-51111-8.

Knowledge, rules, and expert systems

Gallant, S.I. (1995), *Neural Network Learning and Expert Systems,* Cambridge, MA: The MIT Press, ISBN 0-262-07145-2. Chapter headings:: Introduction and Important Definitions; Representation Issues; Perceptron Learning and the Pocket Algorithm; Winner-Take-All Groups or Linear Machines; Autoassociators and One-Shot Learning; Mean Squared Error (MSE) Algorithms; Unsupervised Learning; The Distributed Method and Radial Basis Functions; Computational Learning Theory and the BRD Algorithm; Constructive Algorithms; Backpropagation; Backpropagation: Variations and Applications; Simulated Annealing and Boltzmann Machines; Expert Systems and Neural Networks; Details of the MACIE System; Noise, Redundancy, Fault Detection, and Bayesian Decision Theory; Extracting Rules from Networks; Appendix: Representation Comparisons.

Cloete, I., and Zurada, J.M. (2000), *Knowledge-Based Neurocomputing,* Cambridge, MA: The MIT Press, ISBN 0-262-03274-0. Articles: Knowledge-Based Neurocomputing: Past, Present, and Future; Architectures and Techniques for Knowledge-Based Neurocomputing; Symbolic Knowledge Representation in Recurrent Neural Networks: Insights from Theoretical Models of Computation; A Tutorial on Neurocomputing of Structures; Structural Learning and Rule Discovery; VL[subscript 1]ANN: Transformation of Rules to Artificial Neural Networks; Integrations of Heterogeneous Sources of Partial Domain Knowledge; Approximation of Differential Equations Using Neural Networks; Fynesse: A Hybrid Architecture for Self-Learning Control; Data Mining Techniques for Designing Neural Network Time Series Predictors; Extraction of Decision Trees from Artificial Neural Networks 369; Extraction of Linguistic Rules from Data via Neural Networks and Fuzzy Approximation; Neural Knowledge Processing in Expert Systems.

Learning theory

Wolpert, D.H., ed. (1995) *The Mathematics of Generalization: The Proceedings of the SFI/CNLS Workshop on Formal Approaches to Supervised Learning,* Santa Fe Institute Studies in the Sciences of Complexity, Volume XX, Reading, MA: Addison-Wesley, ISBN: 0201409836.
Articles: The Status of Supervised Learning Science circa 1994 - The Search for a Consensus; Reflections After Refereeing Papers for NIPS; The Probably Approximately Correct (PAC) and Other Learning Models; Decision Theoretic Generalizations of the PAC Model for Neural Net and Other Learning Applications; The Relationship Between PAC, the Statistical Physics Framework, the Bayesian Framework, and the VC Framework; Statistical Physics Models of Supervised Learning; On Exhaustive Learning; A Study of Maximal-Coverage Learning Algorithms; On Bayesian Model Selection; Soft Classification, a.k.a. Risk Estimation, via Penalized Log Likelihood and Smoothing Spline Analysis of Variance; Current Research; Preface to Simplifying Neural Networks by Soft Weight Sharing; Simplifying Neural Networks by Soft Weight Sharing; Error-Correcting Output Codes: A General Method for Improving Multiclass Inductive Learning Programs; Image Segmentation and Recognition.

Anthony, M., and Bartlett, P.L. (1999), *Neural Network Learning: Theoretical Foundations,* Cambridge: Cambridge University Press, ISBN 0-521-57353-X.

Vapnik, V.N. (1998) *Statistical Learning Theory,* NY: Wiley, ISBN: 0471030031
This book is much better than Vapnik's *The Nature of Statistical Learning Theory.*
Chapter headings:
0. Introduction: The Problem of Induction and Statistical Inference;

1. Two Approaches to the Learning Problem;
Appendix: Methods for Solving Ill-Posed Problems;
2. Estimation of the Probability Measure and Problem of Learning;
3. Conditions for Consistency of Empirical Risk Minimization Principle;
4. Bounds on the Risk for Indicator Loss Functions;
Appendix: Lower Bounds on the Risk of the ERM Principle;
5. Bounds on the Risk for Real-Valued Loss Functions;
6. The Structural Risk Minimization Principle;
Appendix: Estimating Functions on the Basis of Indirect Measurements;
7. Stochastic Ill-Posed Problems;
8. Estimating the Values of Functions at Given Points;
9. Perceptrons and Their Generalizations;
10. The Support Vector Method for Estimating Indicator Functions;
11. The Support Vector Method for Estimating Real-Valued Functions;
12. SV Machines for Pattern Recognition; (includes examples of digit recognition)
13. SV Machines for Function Approximations, Regression Estimation, and Signal Processing; (includes an example of positron emission tomography)
14. Necessary and Sufficient Conditions for Uniform Convergence of Frequencies to Their Probabilities;
15. Necessary and Sufficient Conditions for Uniform Convergence of Means to Their Expectations;
16. Necessary and Sufficient Conditions for Uniform One-Sided Convergence of Means to Their Expectations;
Comments and Bibliographical Remarks.

Object oriented programming

The FAQ maintainer is an old-fashioned C programmer and has no expertise in object oriented programming, so he must rely on the readers of comp.ai.neural-nets regarding the merits of books on OOP for NNs.

There are many excellent books about NNs by Timothy Masters (listed elsewhere in the FAQ) that provide C++ code for NNs. If you simply want code that works, these books should satisfy your needs. If you want code that exemplifies the highest standards of object oriented design, you will be disappointed by Masters.

The one book on OOP for NNs that seems to be consistently praised is:

Rogers, Joey (1996), *Object-Oriented Neural Networks in C++,* Academic Press, ISBN 0125931158.
Contents:
1. Introduction
2. Object-Oriented Programming Review
3. Neural-Network Base Classes
4. ADALINE Network
5. Backpropagation Neural Network
6. Self-Organizing Neural Network
7. Bidirectional Associative Memory
Appendix A Support Classes
Appendix B Listings
References and Suggested Reading

However, you will learn very little about NNs other than elementary programming techniques from Rogers. To quote a customer review at the Barnes & Noble web site (http://www.bn.com):

> A reviewer, a scientific programmer, July 19, 2000, **** Long explanation of neural net code - not of neural nets Good OO code for simple 'off the shelf' implementation, very open & fairly extensible for further cusomization. A complete & lucid explanation of the code but pretty weak on the principles, theory, and application of neural networks. Great as a code source, disappointing as a neural network tutorial.

On-line and incremental learning

Saad, D., ed. (1998), *On-Line Learning in Neural Networks,* Cambridge: Cambridge University Press, ISBN 0-521-65263-4.
Articles: Introduction; On-line Learning and Stochastic Approximations; Exact and Perturbation Solutions for the Ensemble Dynamics; A Statistical Study of On-line Learning; On-line Learning in Switching and Drifting Environments with Application to Blind Source Separation; Parameter Adaptation in Stochastic Optimization; Optimal On-line Learning in Multilayer Neural Networks; Universal Asymptotics in Committee Machines with Tree Architecture; Incorporating Curvature Information into On-line Learning; Annealed On-line Learning in Multilayer Neural Networks; On-line Learning of Prototypes and Principal Components; On-line Learning with Time-Correlated Examples; On-line Learning from Finite Training Sets; Dynamics of Supervised Learning with Restricted Training Sets; On-line Learning of a Decision Boundary with and without Queries; A Bayesian Approach to On-line

Learning; Optimal Perceptron Learning: an On-line Bayesian Approach.

Optimization

Cichocki, A. and Unbehauen, R. (1993). *Neural Networks for Optimization and Signal Processing.* NY: John Wiley & Sons, ISBN 0-471-93010-5 (hardbound), 526 pages, $57.95.
Book Webpage (Publisher): http://www.wiley.com/
Additional Information: One has to search.
Chapter headings: Mathematical Preliminaries of Neurocomputing; Architectures and Electronic Implementation of Neural Network Models; Unconstrained Optimization and Learning Algorithms; Neural Networks for Linear, Quadratic Programming and Linear Complementarity Problems; A Neural Network Approach to the On-Line Solution of a System of Linear Algebraic; Equations and Related Problems; Neural Networks for Matrix Algebra Problems; Neural Networks for Continuous, Nonlinear, Constrained Optimization Problems; Neural Networks for Estimation, Identification and Prediction; Neural Networks for Discrete and Combinatorial Optimization Problems.

Pulsed/Spiking networks

Maass, W., and Bishop, C.M., eds. (1999) *Pulsed Neural Networks,* Cambridge, MA: The MIT Press, ISBN: 0262133504.
Articles: Spiking Neurons; Computing with Spiking Neurons; Pulse-Based Computation in VLSI Neural Networks; Encoding Information in Neuronal Activity; Building Silicon Nervous Systems with Dendritic Tree Neuromorphs; A Pulse-Coded Communications Infrastructure; Analog VLSI Pulsed Networks for Perceptive Processing; Preprocessing for Pulsed Neural VLSI Systems; Digital Simulation of Spiking Neural Networks; Populations of Spiking Neurons; Collective Excitation Phenomena and Their Applications; Computing and Learning with Dynamic Synapses; Stochastic Bit-Stream Neural Networks; Hebbian Learning of Pulse Timing in the Barn Owl Auditory System.

Recurrent

Medsker, L.R., and Jain, L.C., eds. (2000), *Recurrent Neural Networks: Design and Applications,* Boca Raton, FL: CRC Press, ISBN 0-8493-7181-3
Articles:
Introduction;
Recurrent Neural Networks for Optimization: The State of the Art;
Efficient Second-Order Learning Algorithms for Discrete-Time Recurrent Neural Networks;
Designing High Order Recurrent Networks for Bayesian Belief Revision;
Equivalence in Knowledge Representation: Automata, Recurrent Neural Networks, and Dynamical Fuzzy Systems;
Learning Long-Term Dependencies in NARX Recurrent Neural Networks;
Oscillation Responses in a Chaotic Recurrent Network;
Lessons from Language Learning;
Recurrent Autoassociative Networks: Developing Distributed Representations of Hierarchically Structured Sequences by Autoassociation;
Comparison of Recurrent Neural Networks for Trajectory Generation;
Training Algorithms for Recurrent Neural Nets that Eliminate the Need for Computation of Error Gradients with Application to Trajectory Production Problem;
Training Recurrent Neural Networks for Filtering and Control;
Remembering How to Behave: Recurrent Neural Networks for Adaptive Robot Behavior

Reinforcement learning

Sutton, R.S., and Barto, A.G. (1998), *Reinforcement Learning: An Introduction,* The MIT Press, ISBN: 0-262193-98-1.
Author's Webpage: http://envy.cs.umass.edu/~rich/sutton.html and http://www-anw.cs.umass.edu/People/barto/barto.html
Book Webpage (Publisher):http://mitpress.mit.edu/book-home.tcl?isbn=0262193981
Additional Information: http://www-anw.cs.umass.edu/~rich/book/the-book.html
Chapter headings: The Problem; Introduction; Evaluative Feedback; The Reinforcement Learning Problem; Elementary Solution Methods; Dynamic Programming; Monte Carlo Methods; Temporal-Difference Learning; A Unified View; Eligibility Traces; Generalization and Function Approximation; Planning and Learning; Dimensions of Reinforcement Learning; Case Studies.

Bertsekas, D. P. and Tsitsiklis, J. N. (1996), *Neuro-Dynamic Programming,* Belmont, MA: Athena Scientific, ISBN 1-886529-10-8.
Author's Webpage: http://www.mit.edu:8001/people/dimitrib/home.html and http://web.mit.edu/jnt/www/home.html
Book Webpage (Publisher):http://world.std.com/~athenasc/ndpbook.html

Speech recognition

Bourlard, H.A., and Morgan, N. (1994), *Connectionist Speech Recognition: A Hybrid Approach,* Boston: Kluwer Academic

Publishers, ISBN: 0792393961.

From The Publisher: Describes the theory and implementation of a method to incorporate neural network approaches into state-of-the-art continuous speech recognition systems based on Hidden Markov Models (HMMs) to improve their performance. In this framework, neural networks (and in particular, multilayer perceptrons or MLPs) have been restricted to well-defined subtasks of the whole system, i.e., HMM emission probability estimation and feature extraction. The book describes a successful five year international collaboration between the authors. The lessons learned form a case study that demonstrates how hybrid systems can be developed to combine neural networks with more traditional statistical approaches. The book illustrates both the advantages and limitations of neural networks in the framework of a statistical system. Using standard databases and comparing with some conventional approaches, it is shown that MLP probability estimation can improve recognition performance. Other approaches are discussed, though there is no such unequivocal experimental result for these methods. Connectionist Speech Recognition: A Hybrid Approach is of use to anyone intending to use neural networks for speech recognition or within the framework provided by an existing successful statistical approach. This includes research and development groups working in the field of speech recognition, both with standard and neural network approaches, as well as other pattern recognition and/or neural network researchers. This book is also suitable as a text for advanced courses on neural networks or speech processing.

Statistics

Cherkassky, V., Friedman, J.H., and Wechsler, H., eds. (1991) *From Statistics to Neural Networks: Theory and Pattern Recognition Applications,* NY: Springer, ISBN 0-387-58199-5.

Kay, J.W., and Titterington, D.M. (1999) *Statistics and Neural Networks: Advances at the Interface,* Oxford: Oxford University Press, ISBN 0-19-852422-6.
Articles: Flexible Discriminant and Mixture Models; Neural Networks for Unsupervised Learning Based on Information Theory; Radial Basis Function Networks and Statistics; Robust Prediction in Many-parameter Models; Density Networks; Latent Variable Models and Data Visualisation; Analysis of Latent Structure Models with Multidimensional Latent Variables; Artificial Neural Networks and Multivariate Statistics.

White, H. (1992b), *Artificial Neural Networks: Approximation and Learning Theory,* Blackwell, ISBN: 1557863296.
Articles: There Exists a Neural Network That Does Not Make Avoidable Mistakes; Multilayer Feedforward Networks Are Universal Approximators; Universal Approximation Using Feedforward Networks with Non-sigmoid Hidden Layer Activation Functions; Approximating and Learning Unknown Mappings Using Multilayer Feedforward Networks with Bounded Weights; Universal Approximation of an Unknown Mapping and Its Derivatives; Neural Network Learning and Statistics; Learning in Artificial Neural Networks: a Statistical Perspective; Some Asymptotic Results for Learning in Single Hidden Layer Feedforward Networks; Connectionist Nonparametric Regression: Multilayer Feedforward Networks Can Learn Arbitrary Mappings; Nonparametric Estimation of Conditional Quantiles Using Neural Networks; On Learning the Derivatives of an Unknown Mapping with Multilayer Feedforward Networks; Consequences and Detection of Misspecified Nonlinear Regression Models; Maximum Likelihood Estimation of Misspecified Models; Some Results for Sieve Estimation with Dependent Observations.

Time-series forecasting

Weigend, A.S. and Gershenfeld, N.A., eds. (1994) *Time Series Prediction: Forecasting the Future and Understanding the Past,* Reading, MA: Addison-Wesley, ISBN 0201626020. Book Webpage (Publisher): http://www2.awl.com/gb/abp/sfi/complexity.html

Unsupervised learning

Kohonen, T. (1995/1997), *Self-Organizing Maps,* 1st ed. 1995, 2nd ed. 1997, Berlin: Springer-Verlag, ISBN 3540620176.

Deco, G. and Obradovic, D. (1996), *An Information-Theoretic Approach to Neural Computing,* NY: Springer-Verlag, ISBN 0-387-94666-7.

Diamantaras, K.I., and Kung, S.Y. (1996) *Principal Component Neural Networks: Theory and Applications,* NY: Wiley, ISBN 0-471-05436-4.

Van Hulle, M.M. (2000), *Faithful Representations and Topographic Maps: From Distortion- to Information-Based Self-Organization,* NY: Wiley, ISBN 0-471-34507-5.

Books for the Beginner

Caudill, M. and Butler, C. (1990). *Naturally Intelligent Systems.* MIT Press: Cambridge, Massachusetts. (ISBN 0-262-03156-6). Book Webpage (Publisher): http://mitpress.mit.edu/book-home.tcl?isbn=0262531135
The authors try to translate mathematical formulas into English. The results are likely to disturb people who appreciate either mathematics or English. Have the authors never heard that "a picture is worth a thousand words"? What few diagrams they have (such as the one on p. 74) tend to be confusing. Their jargon is peculiar even by NN standards; for example, they refer to target values as "mentor inputs" (p. 66). The authors do not understand elementary properties of error functions and optimization

algorithms. For example, in their discussion of the delta rule, the authors seem oblivious to the differences between batch and on-line training, and they attribute magical properties to the algorithm (p. 71):

> [The on-line delta] rule always takes the most efficient route from the current position of the weight vector to the "ideal" position, based on the current input pattern. The delta rule not only minimizes the mean squared error, it does so in the most efficient fashion possible--quite an achievement for such a simple rule.

While the authors realize that backpropagation networks can suffer from local minima, they mistakenly think that counterpropagation has some kind of global optimization ability (p. 202):

> Unlike the backpropagation network, a counterpropagation network cannot be fooled into finding a local minimum solution. This means that the network is guaranteed to find the correct response (or the nearest stored response) to an input, no matter what.

But even though they acknowledge the problem of local minima, the authors are ignorant of the importance of initial weight values (p. 186):

> To teach our imaginary network something using backpropagation, we must start by setting all the adaptive weights on all the neurodes in it to random values. It won't matter what those values are, as long as they are not all the same and not equal to 1.

Like most introductory books, this one neglects the difficulties of getting good generalization--the authors simply declare (p. 8) that "A neural network is able to generalize"!

Chester, M. (1993). *Neural Networks: A Tutorial*, Englewood Cliffs, NJ: PTR Prentice Hall.
Book Webpage (Publisher): http://www.prenhall.com/
Additional Information: Seems to be out of print.
Shallow, sometimes confused, especially with regard to Kohonen networks.

Dayhoff, J. E. (1990). *Neural Network Architectures: An Introduction.* Van Nostrand Reinhold: New York.
Comments from readers of comp.ai.neural-nets: "Like Wasserman's book, Dayhoff's book is also very easy to understand".

Freeman, James (1994). *Simulating Neural Networks with Mathematica,* Addison-Wesley, ISBN: 0-201-56629-X. Book Webpage (Publisher): http://cseng.aw.com/bookdetail.qry?ISBN=0-201-56629-X&ptype=0
Additional Information: Sourcecode available under: ftp://ftp.mathsource.com/pub/Publications/BookSupplements/Freeman-1993
Helps the reader make his own NNs. The mathematica code for the programs in the book is also available through the internet: Send mail to MathSource@wri.com or try http://www.wri.com/ on the World Wide Web.

Freeman, J.A. and Skapura, D.M. (1991). *Neural Networks: Algorithms, Applications, and Programming Techniques,* Reading, MA: Addison-Wesley.
Book Webpage (Publisher): http://www.awl.com/
Additional Information: Seems to be out of print.
A good book for beginning programmers who want to learn how to write NN programs while avoiding any understanding of what NNs do or why they do it.

Gately, E. (1996). *Neural Networks for Financial Forecasting.* New York: John Wiley and Sons, Inc.
Book Webpage (Publisher): http://www.wiley.com/
Additional Information: One has to search.
Franco Insana comments:

```
* Decent book for the neural net beginner
* Very little devoted to statistical framework, although there
    is some formulation of backprop theory
* Some food for thought
* Nothing here for those with any neural net experience
```

McClelland, J. L. and Rumelhart, D. E. (1988). *Explorations in Parallel Distributed Processing: Computational Models of Cognition and Perception* (software manual). The MIT Press.
Book Webpage (Publisher): http://mitpress.mit.edu/book-home.tcl?isbn=026263113X (IBM version) and
http://mitpress.mit.edu/book-home.tcl?isbn=0262631296 (Macintosh)
Comments from readers of comp.ai.neural-nets: "Written in a tutorial style, and includes 2 diskettes of NN simulation programs that can be compiled on MS-DOS or Unix (and they do too !)"; "The programs are pretty reasonable as an introduction to some of the things that NNs can do."; "There are *two* editions of this book. One comes with disks for the IBM PC, the other comes with disks for the Macintosh".

McCord Nelson, M. and Illingworth, W.T. (1990). *A Practical Guide to Neural Nets.* Addison-Wesley Publishing Company, Inc. (ISBN 0-201-52376-0).
Book Webpage (Publisher): http://cseng.aw.com/bookdetail.qry?ISBN=0-201-63378-7&ptype=1174

Lots of applications without technical details, lots of hype, lots of goofs, no formulas.

Muller, B., Reinhardt, J., Strickland, M. T. (1995). *Neural Networks.:An Introduction* (2nd ed.). Berlin, Heidelberg, New York: Springer-Verlag. ISBN 3-540-60207-0. (DOS 3.5" disk included.)
Book Webpage (Publisher): http://www.springer.de/catalog/html-files/deutsch/phys/3540602070.html
Comments from readers of comp.ai.neural-nets: "The book was developed out of a course on neural-network models with computer demonstrations that was taught by the authors to Physics students. The book comes together with a PC-diskette. The book is divided into three parts: (1) Models of Neural Networks; describing several architectures and learing rules, including the mathematics. (2) Statistical Physics of Neural Networks; "hard-core" physics section developing formal theories of stochastic neural networks. (3) Computer Codes; explanation about the demonstration programs. First part gives a nice introduction into neural networks together with the formulas. Together with the demonstration programs a 'feel' for neural networks can be developed."

Orchard, G.A. & Phillips, W.A. (1991). *Neural Computation: A Beginner's Guide.* Lawrence Earlbaum Associates: London.
Comments from readers of comp.ai.neural-nets: "Short user-friendly introduction to the area, with a non-technical flavour. Apparently accompanies a software package, but I haven't seen that yet".

Rao, V.B, and Rao, H.V. (1993). *C++ Neural Networks and Fuzzy Logic.* MIS:Press, ISBN 1-55828-298-x, US $45 incl. disks.
Covers a wider variety of networks than Masters (1993), but is shallow and lacks Masters's insight into practical issues of using NNs.

Wasserman, P. D. (1989). *Neural Computing: Theory & Practice.* Van Nostrand Reinhold: New York. (ISBN 0-442-20743-3)
This is not as bad as some books on NNs. It provides an elementary account of the mechanics of a variety of networks. But it provides no insight into why various methods behave as they do, or under what conditions a method will or will not work well. It has no discussion of efficient training methods such as RPROP or conventional numerical optimization techniques. And, most egregiously, it has no explanation of overfitting and generalization beyond the patently false statement on p. 2 that "It is important to note that the artificial neural network generalizes automatically as a result of its structure"! There is no mention of training, validation, and test sets, or of other methods for estimating generalization error. There is no practical advice on the important issue of choosing the number of hidden units. There is no discussion of early stopping or weight decay. The reader will come away from this book with a grossly oversimplified view of NNs and no concept whatsoever of how to use NNs for practical applications.

Comments from readers of comp.ai.neural-nets: "Wasserman flatly enumerates some common architectures from an engineer's perspective ('how it works') without ever addressing the underlying fundamentals ('why it works') - important basic concepts such as clustering, principal components or gradient descent are not treated. It's also full of errors, and unhelpful diagrams drawn with what appears to be PCB board layout software from the '70s. For anyone who wants to do active research in the field I consider it quite inadequate"; "Okay, but too shallow"; "Quite easy to understand"; "The best bedtime reading for Neural Networks. I have given this book to numerous collegues who want to know NN basics, but who never plan to implement anything. An excellent book to give your manager."

Not-quite-so-introductory Literature

Kung, S.Y. (1993). *Digital Neural Networks,* Prentice Hall, Englewood Cliffs, NJ.

Book Webpage (Publisher): http://www.prenhall.com/books/ptr_0136123260.html
Levine, D. S. (2000). *Introduction to Neural and Cognitive Modeling.* 2nd ed., Lawrence Erlbaum: Hillsdale, N.J.
Comments from readers of comp.ai.neural-nets: "Highly recommended".

Maren, A., Harston, C. and Pap, R., (1990). *Handbook of Neural Computing Applications.* Academic Press. ISBN: 0-12-471260-6. (451 pages)
Comments from readers of comp.ai.neural-nets: "They cover a broad area"; "Introductory with suggested applications implementation".

Pao, Y. H. (1989). *Adaptive Pattern Recognition and Neural Networks* Addison-Wesley Publishing Company, Inc. (ISBN 0-201-12584-6)
Book Webpage (Publisher): http://www.awl.com/
Comments from readers of comp.ai.neural-nets: "An excellent book that ties together classical approaches to pattern recognition with Neural Nets. Most other NN books do not even mention conventional approaches."

Refenes, A. (Ed.) (1995). *Neural Networks in the Capital Markets.* Chichester, England: John Wiley and Sons, Inc.
Book Webpage (Publisher): http://www.wiley.com/
Additional Information: One has to search.
Franco Insana comments:

```
* Not for the beginner
* Excellent introductory material presented by editor in first 5
  chapters, which could be a valuable reference source for any
  practitioner
* Very thought-provoking
* Mostly backprop-related
```

```
* Most contributors lay good statistical foundation
* Overall, a wealth of information and ideas, but the reader has to
  sift through it all to come away with anything useful
```

Simpson, P. K. (1990). *Artificial Neural Systems: Foundations, Paradigms, Applications and Implementations.* Pergamon Press: New York.
Comments from readers of comp.ai.neural-nets: "Contains a very useful 37 page bibliography. A large number of paradigms are presented. On the negative side the book is very shallow. Best used as a complement to other books".

Wasserman, P.D. (1993). *Advanced Methods in Neural Computing.* Van Nostrand Reinhold: New York (ISBN: 0-442-00461-3).
Comments from readers of comp.ai.neural-nets: "Several neural network topics are discussed e.g. Probalistic Neural Networks, Backpropagation and beyond, neural control, Radial Basis Function Networks, Neural Engineering. Furthermore, several subjects related to neural networks are mentioned e.g. genetic algorithms, fuzzy logic, chaos. Just the functionality of these subjects is described; enough to get you started. Lots of references are given to more elaborate descriptions. Easy to read, no extensive mathematical background necessary."

Zeidenberg. M. (1990). *Neural Networks in Artificial Intelligence.* Ellis Horwood, Ltd., Chichester.
Comments from readers of comp.ai.neural-nets: "Gives the AI point of view".

Zornetzer, S. F., Davis, J. L. and Lau, C. (1990). *An Introduction to Neural and Electronic Networks.* Academic Press. (ISBN 0-12-781881-2)
Comments from readers of comp.ai.neural-nets: "Covers quite a broad range of topics (collection of articles/papers)."; "Provides a primer-like introduction and overview for a broad audience, and employs a strong interdisciplinary emphasis".

Zurada, Jacek M. (1992). *Introduction To Artificial Neural Systems.* Hardcover, 785 Pages, 317 Figures, ISBN 0-534-95460-X, 1992, PWS Publishing Company, Price: $56.75 (includes shipping, handling, and the ANS software diskette). Solutions Manual available.
Comments from readers of comp.ai.neural-nets: "Cohesive and comprehensive book on neural nets; as an engineering-oriented introduction, but also as a research foundation. Thorough exposition of fundamentals, theory and applications. Training and recall algorithms appear in boxes showing steps of algorithms, thus making programming of learning paradigms easy. Many illustrations and intuitive examples. Winner among NN textbooks at a senior UG/first year graduate level-[175 problems]." Contents: Intro, Fundamentals of Learning, Single-Layer & Multilayer Perceptron NN, Assoc. Memories, Self-organizing and Matching Nets, Applications, Implementations, Appendix)

Books with Source Code (C, C++)

Blum, Adam (1992), *Neural Networks in C++,* Wiley.

Review by Ian Cresswell. (For a review of the text, see "The Worst" below.)

Mr Blum has not only contributed a masterpiece of NN inaccuracy but also seems to lack a fundamental understanding of Object Orientation.

The excessive use of virtual methods (see page 32 for example), the inclusion of unnecessary 'friend' relationships (page 133) and a penchant for operator overloading (pick a page!) demonstrate inability in C++ and/or OO.

The introduction to OO that is provided trivialises the area and demonstrates a distinct lack of direction and/or understanding.

The public interfaces to classes are overspecified and the design relies upon the flawed neuron/layer/network model.

There is a notable disregard for any notion of a robust class hierarchy which is demonstrated by an almost total lack of concern for inheritance and associated reuse strategies.

The attempt to rationalise differing types of Neural Network into a single very shallow but wide class hierarchy is naive.

The general use of the 'float' data type would cause serious hassle if this software could possibly be extended to use some of the more sensitive variants of backprop on more difficult problems. It is a matter of great fortune that such software is unlikely to be reusable and will therefore, like all good dinosaurs, disappear with the passage of time.

The irony is that there is a card in the back of the book asking the unfortunate reader to part with a further $39.95 for a copy of the software (already included in print) on a 5.25" disk.

The author claims that his work provides an 'Object Oriented Framework ...'. This can best be put in his own terms (Page 137):

```
... garble(float noise) ...
```

Swingler, K. (1996), *Applying Neural Networks: A Practical Guide,* London: Academic Press.

Review by Ian Cresswell. (For a review of the text, see "The Worst" below.)

Before attempting to review the code associated with this book it should be clearly stated that it is supplied as an extra--almost as an afterthought. This may be a wise move.

Although not as bad as other (even commercial) implementations, the code provided lacks proper OO structure and is typical of C++ written in a C style.

Style criticisms include:

1. The use of public data fields within classes (loss of encapsulation).
2. Classes with no protected or private sections.
3. Little or no use of inheritance and/or run-time polymorphism.
4. Use of floats not doubles (a common mistake) to store values for connection weights.
5. Overuse of classes and public methods. The network class has 59 methods in its public section.
6. Lack of planning is evident for the construction of a class hierarchy.

This code is without doubt written by a rushed C programmer. Whilst it would require a C++ compiler to be successfully used, it lacks the tight (optimised) nature of good C and the high level of abstraction of good C++.

In a generous sense the code is free and the author doesn't claim any expertise in software engineering. It works in a limited sense but would be difficult to extend and/or reuse. It's fine for demonstration purposes in a stand-alone manner and for use with the book concerned.

If you're serious about nets you'll end up rewriting the whole lot (or getting something better).

The Worst

How not to use neural nets in any programming language

> Blum, Adam (1992), *Neural Networks in C++*, NY: Wiley.

> Welstead, Stephen T. (1994), *Neural Network and Fuzzy Logic Applications in C/C++*, NY: Wiley.

(For a review of Blum's source code, see "Books with Source Code" above.)

Both Blum and Welstead contribute to the dangerous myth that any idiot can use a neural net by dumping in whatever data are handy and letting it train for a few days. They both have little or no discussion of generalization, validation, and overfitting. Neither provides any valid advice on choosing the number of hidden nodes. If you have ever wondered where these stupid "rules of thumb" that pop up frequently come from, here's a source for one of them:

> "A rule of thumb is for the size of this [hidden] layer to be somewhere between the input layer size ... and the output layer size ..." Blum, p. 60.

(John Lazzaro tells me he recently "reviewed a paper that cited this rule of thumb--and referenced this book! Needless to say, the final version of that paper didn't include the reference!")

Blum offers some profound advice on choosing inputs:

> "The next step is to pick as many input factors as possible that might be related to [the target]."

Blum also shows a deep understanding of statistics:

> "A statistical model is simply a more indirect way of learning correlations. With a neural net approach, we model the problem directly." p. 8.

Blum at least mentions some important issues, however simplistic his advice may be. Welstead just ignores them. What Welstead gives you is code--vast amounts of code. I have no idea how anyone could write *that* much code for a simple feedforward NN. Welstead's approach to validation, in his chapter on financial forecasting, is to reserve *two* cases for the validation set!

My comments apply only to the text of the above books. I have not examined or attempted to compile the code.

An impractical guide to neural nets

> Swingler, K. (1996), *Applying Neural Networks: A Practical Guide,* London: Academic Press.

(For a review of the source code, see "Books with Source Code" above.)

This book has lots of good advice liberally sprinkled with errors, incorrect formulas, some bad advice, and some very serious mistakes. Experts will learn nothing, while beginners will be unable to separate the useful information from the dangerous. For example, there is a chapter on "Data encoding and re-coding" that would be very useful to beginners if it were accurate, but the formula for the standard deviation is wrong, and the description of the softmax function is of something entirely different than softmax (see What is a softmax activation function?). Even more dangerous is the statement on p. 28 that "Any pair of variables with high covariance are dependent, and one may be chosen to be discarded." Although high correlations can be used to identify redundant inputs, it is incorrect to use high covariances for this purpose, since a covariance can be high simply because one of the inputs has a high standard deviation.

The most ludicrous thing I've found in the book is the claim that Hecht-Neilsen used Kolmogorov's theorem to show that "you will never require more than twice the number of hidden units as you have inputs" (p. 53) in an MLP with one hidden layer. Actually, Hecht-Neilsen, says "the direct usefulness of this result is doubtful, because no constructive method for developing the [output activation] functions is known." Then Swingler implies that V. Kurkova (1991, "Kolmogorov's theorem is relevant," Neural Computation, 3, 617-622) confirmed this alleged upper bound on the number of hidden units, saying that, "Kurkova was able to restate Kolmogorov's theorem in terms of a set of sigmoidal functions." If Kolmogorov's theorem, or Hecht-Neilsen's adaptation of it, could be restated in terms of known sigmoid activation functions in the (single) hidden and output layers, then Swingler's alleged upper bound would be correct, but in fact no such restatement of Kolmogorov's theorem is possible, and Kurkova did not claim to prove any such restatement. Swingler omits the crucial details that Kurkova used two hidden layers, staircase-like activation functions (not ordinary sigmoidal functions such as the logistic) in the first hidden layer, and a potentially large number of units in the second hidden layer. Kurkova later estimated the number of units required for uniform approximation within an error $epsilon$ as $nm(m+1)$ in the first hidden layer and $m^2(m+1)^n$ in the second hidden layer, where n is the number of inputs and m "depends on $epsilon/||f||$ as well as on the rate with which f increases distances." In other words, Kurkova says nothing to support Swinglers advice (repeated on p. 55), "Never choose h to be more than twice the number of input units." Furthermore, constructing a counter example to Swingler's advice is trivial: use one input and one output, where the output is the sine of the input, and the domain of the input extends over many cycles of the sine wave; it is obvious that many more than two hidden units are required. For some sound information on choosing the number of hidden units, see How many hidden units should I use?

Choosing the number of hidden units is one important aspect of getting good generalization, which is the most crucial issue in neural network training. There are many other considerations involved in getting good generalization, and Swingler makes several more mistakes in this area:

- There is dangerous misinformation on p. 55, where Swingler says, "If a data set contains no noise, then there is no risk of overfitting as there is nothing to overfit." It is true that overfitting is more common with noisy data, but severe overfitting can occur with noise-free data, even when there are more training cases than weights. There is an example of such overfitting under How many hidden layers should I use?

- Regarding the use of added noise (jitter) in training, Swingler says on p. 60, "The more noise you add, the more general your model becomes." This statement makes no sense as it stands (it would make more sense if "general" were changed to "smooth"), but it could certainly encourage a beginner to use far too much jitter--see What is jitter? (Training with noise).

- On p. 109, Swingler describes leave-one-out cross-validation, which he ascribes to Hecht-Neilsen. But Swingler concludes, "the method provides you with L minus 1 networks to choose from; none of which has been validated properly," completely missing the point that cross-validation provides an estimate of the generalization error of a network trained on the entire training set of L cases--see What are cross-validation and bootstrapping? Also, there are L leave-one-out networks, not L-1.

While Swingler has some knowldge of statistics, his expertise is not sufficient for him to detect that certain articles on neural nets are statistically nonsense. For example, on pp. 139-140 he uncritically reports a method that allegedly obtains error bars by doing a simple linear regression on the target vs. output scores. To a trained statistician, this method is obviously wrong (and, as usual in this book, the formula for variance given for this method on p. 150 is wrong). On p. 110, Swingler reports an article that attempts to apply bootstrapping to neural nets, but this article is also obviously wrong to anyone familiar with bootstrapping. While Swingler cannot be blamed entirely for accepting these articles at face value, such misinformation provides yet more hazards for beginners.

Swingler addresses many important practical issues, and often provides good practical advice. But the peculiar combination of much good advice with some extremely bad advice, a few examples of which are provided above, could easily seduce a beginner into thinking that the book as a whole is reliable. It is this danger that earns the book a place in "The Worst" list.

Bad science writing

Dewdney, A.K. (1997), *Yes, We Have No Neutrons: An Eye-Opening Tour through the Twists and Turns of Bad Science,* NY: Wiley.

This book, allegedly an expose of bad science, contains only one chapter of 19 pages on "the neural net debacle" (p. 97). Yet this chapter is so egregiously misleading that the book has earned a place on "The Worst" list. A detailed criticism of this chapter, along with some other sections of the book, can be found at ftp://ftp.sas.com/pub/neural/badscience.html. Other chapters of the book are reviewed in the November, 1997, issue of *Scientific American.*

Subject: Journals and magazines about Neural Networks?

```
[to be added: comments on speed of reviewing and publishing,
               whether they accept TeX format or ASCII by e-mail, etc.]
```

A. Dedicated Neural Network Journals:

```
Title:   Neural Networks
Publish: Pergamon Press
Address: Pergamon Journals Inc., Fairview Park, Elmsford,
         New York 10523, USA and Pergamon Journals Ltd.
         Headington Hill Hall, Oxford OX3, 0BW, England
Freq.:   10 issues/year (vol. 1 in 1988)
Cost/Yr: Free with INNS or JNNS or ENNS membership ($45?),
         Individual $65, Institution $175
ISSN #:  0893-6080
URL:     http://www.elsevier.nl/locate/inca/841
Remark:  Official Journal of International Neural Network Society (INNS),
         European Neural Network Society (ENNS) and Japanese Neural
         Network Society (JNNS).
         Contains Original Contributions, Invited Review Articles, Letters
         to Editor, Book Reviews, Editorials, Announcements, Software Surveys.


Title:   Neural Computation
Publish: MIT Press
Address: MIT Press Journals, 55 Hayward Street Cambridge,
         MA 02142-9949, USA, Phone: (617) 253-2889
Freq.:   Quarterly (vol. 1 in 1989)
Cost/Yr: Individual $45, Institution $90, Students $35; Add $9 Outside USA
ISSN #:  0899-7667
URL:     http://mitpress.mit.edu/journals-legacy.tcl
Remark:  Combination of Reviews (10,000 words), Views (4,000 words)
         and Letters (2,000 words).  I have found this journal to be of
         outstanding quality.
         (Note: Remarks supplied by Mike Plonski "plonski@aero.org")


Title:   NEURAL COMPUTING SURVEYS
Publish: Lawrence Erlbaum Associates
Address: 10 Industrial Avenue, Mahwah, NJ  07430-2262, USA
Freq.:   Yearly
Cost/Yr: Free on-line
ISSN #:  1093-7609
URL:     http://www.icsi.berkeley.edu/~jagota/NCS/
Remark:  One way to cope with the exponential increase in the number
         of articles published in recent years is to ignore most of
         them. A second, perhaps more satisfying, approach is to
         provide a forum that encourages the regular production --
         and perusal -- of high-quality survey articles. This is
         especially useful in an inter-disciplinary, evolving field
         such as neural networks. This journal aims to bring the
         second view-point to bear. It is intended to

         * encourage researchers to write good survey papers.
         * motivate researchers to look here first to check
           what's known on an unfamiliar topic.


Title:   IEEE Transactions on Neural Networks
Publish: Institute of Electrical and Electronics Engineers (IEEE)
Address: IEEE Service Cemter, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ,
         08855-1331 USA. Tel: (201) 981-0060
Cost/Yr: $10 for Members belonging to participating IEEE societies
Freq.:   Quarterly (vol. 1 in March 1990)
URL:     http://www.ieee.org/nnc/pubs/transactions.html
Remark:  Devoted to the science and technology of neural networks
         which disclose significant  technical knowledge, exploratory
         developments and applications of neural networks from biology to
         software to hardware.  Emphasis is on artificial neural networks.
         Specific aspects include self organizing systems, neurobiological
         connections, network dynamics and architecture, speech recognition,
         electronic and photonic implementation, robotics and controls.
         Includes Letters concerning new research results.
         (Note: Remarks are from journal announcement)


Title:   IEEE Transactions on Evolutionary Computation
Publish: Institute of Electrical and Electronics Engineers (IEEE)
Address: IEEE Service Cemter, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ,
         08855-1331 USA. Tel: (201) 981-0060
Cost/Yr: $10 for Members belonging to participating IEEE societies
Freq.:   Quarterly (vol. 1 in May 1997)
```

```
URL:     http://engine.ieee.org/nnc/pubs/transactions.html
Remark:  The IEEE Transactions on Evolutionary Computation will publish archival
         journal quality original papers in evolutionary computation and related
         areas, with particular emphasis on the practical application of the
         techniques to solving real problems in industry, medicine, and other
         disciplines.  Specific techniques include but are not limited to
         evolution strategies, evolutionary programming, genetic algorithms, and
         associated methods of genetic programming and classifier systems.  Papers
         emphasizing mathematical results should ideally seek to put these results
         in the context of algorithm design, however purely theoretical papers will
         be considered.  Other papers in the areas of cultural algorithms, artificial
         life, molecular computing, evolvable hardware, and the use of simulated
         evolution to gain a better understanding of naturally evolved systems are
         also encouraged.
         (Note: Remarks are from journal CFP)


Title:   International Journal of Neural Systems
Publish: World Scientific Publishing
Address: USA: World Scientific Publishing Co., 1060 Main Street, River Edge,
         NJ 07666. Tel: (201) 487 9655; Europe: World Scientific Publishing
         Co. Ltd., 57 Shelton Street, London WC2H 9HE, England.
         Tel: (0171) 836 0888; Asia: World Scientific Publishing Co. Pte. Ltd.,
         1022 Hougang Avenue 1 #05-3520, Singapore 1953, Rep. of Singapore
         Tel: 382 5663.
Freq.:   Quarterly (Vol. 1 in 1990)
Cost/Yr: Individual $122, Institution $255 (plus $15-$25 for postage)
ISSN #:  0129-0657 (IJNS)
Remark:  The International Journal of Neural Systems is a quarterly
         journal which covers information processing in natural
         and artificial neural systems. Contributions include research papers,
         reviews, and Letters to the Editor - communications under 3,000
         words in length, which are published within six months of receipt.
         Other contributions are typically published within nine months.
         The journal presents a fresh undogmatic attitude towards this
         multidisciplinary field and aims to be a forum for novel ideas and
         improved understanding of collective and cooperative phenomena with
         computational capabilities.
         Papers should be submitted to World Scientific's UK office. Once a
         paper is accepted for publication, authors are invited to e-mail
         the LaTeX source file of their paper in order to expedite publication.


Title:   International Journal of Neurocomputing
Publish: Elsevier Science Publishers, Journal Dept.; PO Box 211;
         1000 AE Amsterdam, The Netherlands
Freq.:   Quarterly (vol. 1 in 1989)
URL:     http://www.elsevier.nl/locate/inca/505628


Title:   Neural Processing Letters
Publish: Kluwer Academic publishers
Address: P.O. Box 322, 3300 AH Dordrecht, The Netherlands
Freq:    6 issues/year (vol. 1 in 1994)
Cost/Yr: Individuals $198, Institution $400 (including postage)
ISSN #:  1370-4621
URL:     http://www.wkap.nl/journalhome.htm/1370-4621
Remark:  The aim of the journal is to rapidly publish new ideas, original
         developments and work in progress.  Neural Processing Letters
         covers all aspects of the Artificial Neural Networks field.
         Publication delay is about 3 months.


Title:   Neural Network News
Publish: AIWeek Inc.
Address: Neural Network News, 2555 Cumberland Parkway, Suite 299,
         Atlanta, GA 30339 USA. Tel: (404) 434-2187
Freq.:   Monthly (beginning September 1989)
Cost/Yr: USA and Canada $249, Elsewhere $299
Remark:  Commercial Newsletter


Title:   Network: Computation in Neural Systems
Publish: IOP Publishing Ltd
Address: Europe: IOP Publishing Ltd, Techno House, Redcliffe Way, Bristol
         BS1 6NX, UK; IN USA: American Institute of Physics, Subscriber
         Services 500 Sunnyside Blvd., Woodbury, NY  11797-2999
Freq.:   Quarterly (1st issue 1990)
Cost/Yr: USA: $180,  Europe: 110 pounds
URL:     http://www.iop.org/Journals/ne
Remark:  Description: "a forum for integrating theoretical and experimental
         findings across relevant interdisciplinary boundaries."  Contents:
         Submitted articles reviewed by two technical referees  paper's
         interdisciplinary format and accessability."  Also Viewpoints and
         Reviews commissioned by the editors, abstracts (with reviews) of
         articles published in other journals, and book reviews.
         Comment: While the price discourages me (my comments are based
         upon a free sample copy), I think that the journal succeeds
```

```
                 very well.  The highest density of interesting articles I
                 have found in any journal.
                 (Note: Remarks supplied by kehoe@csufres.CSUFresno.EDU)

Title:     Connection Science: Journal of Neural Computing,
           Artificial Intelligence and Cognitive Research
Publish: Carfax Publishing
Address: Europe: Carfax Publishing Company, PO Box 25, Abingdon, Oxfordshire
           OX14 3UE, UK.
           USA: Carfax Publishing Company, PO Box 2025, Dunnellon, Florida
           34430-2025, USA
           Australia: Carfax Publishing Company, Locked Bag 25, Deakin,
           ACT 2600, Australia
Freq.:   Quarterly (vol. 1 in 1989)
Cost/Yr: Personal rate:
           48 pounds (EC) 66 pounds (outside EC) US$118 (USA and Canada)
           Institutional rate:
           176 pounds (EC) 198 pounds (outside EC) US$340 (USA and Canada)

Title:     International Journal of Neural Networks
Publish: Learned Information
Freq.:   Quarterly (vol. 1 in 1989)
Cost/Yr: 90 pounds
ISSN #:  0954-9889
Remark:  The journal contains articles, a conference report (at least the
           issue I have), news and a calendar.
           (Note: remark provided by J.R.M. Smits "anjos@sci.kun.nl")

Title:     Sixth Generation Systems (formerly Neurocomputers)
Publish: Gallifrey Publishing
Address: Gallifrey Publishing, PO Box 155, Vicksburg, Michigan, 49097, USA
           Tel: (616) 649-3772, 649-3592 fax
Freq.    Monthly (1st issue January, 1987)
ISSN #:  0893-1585
Editor:  Derek F. Stubbs
Cost/Yr: $79 (USA, Canada), US$95 (elsewhere)
Remark:  Runs eight to 16 pages monthly. In 1995 will go to floppy disc-based
publishing with databases +, "the equivalent to 50 pages per issue are
planned." Often focuses on specific topics: e.g., August, 1994 contains two
articles: "Economics, Times Series and the Market," and "Finite Particle
Analysis - [part] II."  Stubbs also directs the company Advanced Forecasting
Technologies. (Remark by Ed Rosenfeld: ier@aol.com)

Title:    JNNS Newsletter (Newsletter of the Japan Neural Network Society)
Publish: The Japan Neural Network Society
Freq.:   Quarterly (vol. 1 in 1989)
Remark:  (IN JAPANESE LANGUAGE) Official Newsletter of the Japan Neural
           Network Society(JNNS)
           (Note: remarks by Osamu Saito "saito@nttica.NTT.JP")

Title:    Neural Networks Today
Remark:  I found this title in a bulletin board of october last year.
           It was a message of Tim Pattison, timpatt@augean.OZ
           (Note: remark provided by J.R.M. Smits "anjos@sci.kun.nl")

Title:    Computer Simulations in Brain Science

Title:    Internation Journal of Neuroscience

Title:    Neural Network Computation
Remark:  Possibly the same as "Neural Computation"

Title:    Neural Computing and Applications
Freq.:   Quarterly
Publish: Springer Verlag
Cost/yr: 120 Pounds
Remark:  Is the journal of the Neural Computing Applications Forum.
           Publishes original research and other information
           in the field of practical applications of neural computing.
```

B. NN Related Journals:

```
Title:    Biological Cybernetics (Kybernetik)
Publish: Springer Verlag
Remark:  Monthly (vol. 1 in 1961)

Title:    Various IEEE Transactions and Magazines
Publish: IEEE
Remark:  Primarily see IEEE Trans. on System, Man and Cybernetics;
           Various Special Issues: April 1990 IEEE Control Systems
           Magazine.; May 1989 IEEE Trans. Circuits and Systems.;
           July 1988 IEEE Trans. Acoust. Speech Signal Process.
```

```
Title:   The Journal of Experimental and Theoretical Artificial Intelligence
Publish: Taylor & Francis, Ltd.
Address: London, New York, Philadelphia
Freq.:   ? (1st issue Jan 1989)
Remark:  For submission information, please contact either of the editors:
         Eric Dietrich                    Chris Fields
         PACSS - Department of Philosophy    Box 30001/3CRL
         SUNY Binghamton                  New Mexico State University
         Binghamton, NY 13901             Las Cruces, NM 88003-0001
         dietrich@bingvaxu.cc.binghamton.edu  cfields@nmsu.edu


Title:   The Behavioral and Brain Sciences
Publish: Cambridge University Press
Remark:  (Remarks by Don Wunsch
         This is a delightful journal that encourages discussion on a
         variety of controversial topics.  I have especially enjoyed
         reading some papers in there by Dana Ballard and Stephen
         Grossberg (separate papers, not collaborations) a few years
         back.  They have a really neat concept: they get a paper,
         then invite a number of noted scientists in the field to
         praise it or trash it.  They print these commentaries, and
         give the author(s) a chance to make a rebuttal or
         concurrence.  Sometimes, as I'm sure you can imagine, things
         get pretty lively. Their reviewers are called something like
         Behavioral and Brain Associates, and I believe they have to
         be nominated by current associates, and should be fairly
         well established in the field. The main thing is that I liked
         the articles I read.


Title:   International Journal of Applied Intelligence
Publish: Kluwer Academic Publishers
Remark:  first issue in 1990(?)


Title:   International Journal of Modern Physics C
Publish: USA: World Scientific Publishing Co., 1060 Main Street, River Edge,
         NJ 07666. Tel: (201) 487 9655; Europe: World Scientific Publishing
         Co. Ltd., 57 Shelton Street, London WC2H 9HE, England.
         Tel: (0171) 836 0888; Asia: World Scientific Publishing Co. Pte. Ltd.,
         1022 Hougang Avenue 1 #05-3520, Singapore 1953, Rep. of Singapore
         Tel: 382 5663.
Freq:    bi-monthly
Eds:     H. Herrmann, R. Brower, G.C. Fox and S Nose


Title:   Machine Learning
Publish: Kluwer Academic Publishers
Address: Kluwer Academic Publishers
         P.O. Box 358
         Accord Station
         Hingham, MA 02018-0358 USA
Freq.:   Monthly (8 issues per year; increasing to 12 in 1993)
Cost/Yr: Individual $140 (1992); Member of AAAI or CSCSI $88
Remark:  Description: Machine Learning is an international forum for
         research on computational approaches to learning.  The journal
         publishes articles reporting substantive research results on a
         wide range of learning methods applied to a variety of task
         domains.  The ideal paper will make a theoretical contribution
         supported by a computer implementation.
         The journal has published many key papers in learning theory,
         reinforcement learning, and decision tree methods.  Recently
         it has published a special issue on connectionist approaches
         to symbolic reasoning.  The journal regularly publishes
         issues devoted to genetic algorithms as well.


Title:   INTELLIGENCE - The Future of Computing
Published by: Intelligence
Address: INTELLIGENCE, P.O. Box 20008, New York, NY 10025-1510, USA,
212-222-1123 voice & fax; email: ier@aol.com, CIS: 72400,1013
Freq.   Monthly plus four special reports each year (1st issue: May, 1984)
ISSN #: 1042-4296
Editor: Edward Rosenfeld
Cost/Yr: $395 (USA), US$450 (elsewhere)
Remark:  Has absorbed several other newsletters, like Synapse/Connection
         and Critical Technology Trends (formerly AI Trends).
         Covers NN, genetic algorithms, fuzzy systems, wavelets, chaos
         and other advanced computing approaches, as well as molecular
         computing and nanotechnology.


Title:   Journal of Physics A: Mathematical and General
Publish: Inst. of Physics, Bristol
Freq:    24 issues per year.
Remark:  Statistical mechanics aspects of neural networks
         (mostly Hopfield models).
```

```
Title:   Physical Review A: Atomic, Molecular and Optical Physics
Publish: The American Physical Society (Am. Inst. of Physics)
Freq:    Monthly
Remark:  Statistical mechanics of neural networks.

Title:   Information Sciences
Publish: North Holland (Elsevier Science)
Freq.:   Monthly
ISSN:    0020-0255
Editor:  Paul P. Wang; Department of Electrical Engineering; Duke University;
         Durham, NC 27706, USA
```

Subject: Conferences and Workshops on Neural Networks?

- The journal "Neural Networks" has a list of conferences, workshops and meetings in each issue.
- NEuroNet maintains a list of Neural Network Events at http://www.kcl.ac.uk/neuronet/events/index.html
- The IEEE Neural Network Council maintains a list of conferences at http://www.ieee.org/nnc.
- Conferences, workshops, and other events concerned with neural networks, inductive learning, genetic algorithms, data mining, agents, applications of AI, pattern recognition, vision, and related fields. are listed at Georg Thimm's web page http://www.drc.ntu.edu.sg/users/mgeorg/enter.epl

----

Subject: Neural Network Associations?

1. **International Neural Network Society (INNS).**

   INNS membership includes subscription to "Neural Networks", the official journal of the society. Membership is $55 for non-students and $45 for students per year. Address: INNS Membership, P.O. Box 491166, Ft. Washington, MD 20749.

2. **International Student Society for Neural Networks (ISSNNets).**

   Membership is $5 per year. Address: ISSNNet, Inc., P.O. Box 15661, Boston, MA 02215 USA

3. **Women In Neural Network Research and technology (WINNERS).**

   Address: WINNERS, c/o Judith Dayhoff, 11141 Georgia Ave., Suite 206, Wheaton, MD 20902. Phone: 301-933-9000.

4. **European Neural Network Society (ENNS)**

   ENNS membership includes subscription to "Neural Networks", the official journal of the society. Membership is currently (1994) 50 UK pounds (35 UK pounds for students) per year. Address: ENNS Membership, Centre for Neural Networks, King's College London, Strand, London WC2R 2LS, United Kingdom.

5. **Japanese Neural Network Society (JNNS)**

   Address: Japanese Neural Network Society; Department of Engineering, Tamagawa University; 6-1-1, Tamagawa Gakuen, Machida City, Tokyo; 194 JAPAN; Phone: +81 427 28 3457, Fax: +81 427 28 3597

6. **Association des Connexionnistes en THese (ACTH)**

   (the French Student Association for Neural Networks); Membership is 100 FF per year; Activities: newsletter, conference (every year), list of members, electronic forum; Journal 'Valgo' (ISSN 1243-4825); WWW page: http://www.supelec-rennes.fr /acth/welcome.html ; Contact: acth@loria.fr

7. **Neurosciences et Sciences de l'Ingenieur (NSI)**

   Biology & Computer Science Activity : conference (every year) Address : NSI - TIRF / INPG 46 avenue Felix Viallet 38031 Grenoble Cedex FRANCE

8. **IEEE Neural Networks Council**

   Web page at http://www.ieee.org/nnc

9. **SNN (Foundation for Neural Networks)**

The Foundation for Neural Networks (SNN) is a university based non-profit organization that stimulates basic and applied research on neural networks in the Netherlands. Every year SNN orgines a symposium on Neural Networks. See http://www.mbfys.kun.nl/SNN/.

You can find nice lists of NN societies in the WWW at http://www.emsl.pnl.gov:2080/proj/neuron/neural/societies.html and at http://www.ieee.org:80/nnc/research/othernnsoc.html.

Subject: Mailing lists, BBS, CD-ROM?

See also "Other NN links?" in Part 7 of the FAQ.

1. **Machine Learning mailing list**

   http://groups.yahoo.com/group/machine-learning/

   The Machine Learning mailing list is an unmoderated mailing list intended for people in Computer Sciences, Statistics, Mathematics, and other areas or disciplines with interests in Machine Learning. Researchers, practitioners, and users of Machine Learning in academia, industry, and government are encouraged to join the list to discuss and exchange ideas regarding any aspect of Machine Learning, e.g., various learning algorithms, data pre-processing, variable selection mechanism, instance selection, and applications to real-world problems.

   You can post, read, and reply messages on the Web. Or you can choose to receive messages as individual emails, daily summaries, daily full-text digest, or read them on the Web only.

2. **The Connectionists Mailing List**

   http://www.cnbc.cmu.edu/other/connectionists.html

   CONNECTIONISTS is a moderated mailing list for discussion of technical issues relating to neural computation, and dissemination of professional announcements such as calls for papers, book announcements, and electronic preprints. CONNECTIONISTS is focused on meeting the needs of active researchers in the field, not on answering questions from beginners.

3. **Central Neural System Electronic Bulletin Board**

   ```
   URL: ftp://www.centralneuralsystem.com/pub/CNS/bbs
   Supported by: Wesley R. Elsberry
                 3027 Macaulay Street
                 San Diego, CA 92106
   Email: welsberr@inia.cls.org
   Alternative URL: http://www.cs.cmu.edu/afs/cs.cmu.edu/project/ai-repository/ai/areas/neural/cns/0.html
   ```

   Many MS-DOS PD and shareware simulations, source code, benchmarks, demonstration packages, information files; some Unix, Macintosh, Amiga related files. Also available are files on AI, AI Expert listings 1986-1991, fuzzy logic, genetic algorithms, artificial life, evolutionary biology, and many Project Gutenberg and Wiretap e-texts.

4. **AI CD-ROM**

   Network Cybernetics Corporation produces the "AI CD-ROM". It is an ISO-9660 format CD-ROM and contains a large assortment of software related to artificial intelligence, artificial life, virtual reality, and other topics. Programs for OS/2, MS-DOS, Macintosh, UNIX, and other operating systems are included. Research papers, tutorials, and other text files are included in ASCII, RTF, and other universal formats. The files have been collected from AI bulletin boards, Internet archive sites, University computer deptartments, and other government and civilian AI research organizations. Network Cybernetics Corporation intends to release annual revisions to the AI CD-ROM to keep it up to date with current developments in the field. The AI CD-ROM includes collections of files that address many specific AI/AL topics including Neural Networks (Source code and executables for many different platforms including Unix, DOS, and Macintosh. ANN development tools, example networks, sample data, tutorials. A complete collection of Neural Digest is included as well.) The AI CD-ROM may be ordered directly by check, money order, bank draft, or credit card from:

   ```
   Network Cybernetics Corporation;
   4201 Wingren Road Suite 202;
   Irving, TX 75062-2763;
   Tel 214/650-2002;
   ```

   Fax 214/650-1929;

The cost is $129 per disc + shipping ($5/disc domestic or $10/disc foreign) (See the comp.ai FAQ for further details)

Subject: How to benchmark learning methods?

The NN benchmarking resources page at http://wwwipd.ira.uka.de/~prechelt/NIPS_bench.html was created after a NIPS 1995 workshop on NN benchmarking. The page contains pointers to various papers on proper benchmarking methodology and to various sources of datasets.

Benchmark studies require some familiarity with the statistical design and analysis of experiments. There are many textbooks on this subject, of which Cohen (1995) will probably be of particular interest to researchers in neural nets and machine learning (see also the review of Cohen's book by Ron Kohavi in the International Journal of Neural Systems, which can be found on-line at http://robotics.stanford.edu/users/ronnyk/ronnyk-bib.html).

Reference:

   Cohen, P.R. (1995), *Empirical Methods for Artificial Intelligence,* Cambridge, MA: The MIT Press.

Subject: Databases for experimentation with NNs?

1. **UCI machine learning database**

   A large collection of data sets accessible via anonymous FTP at ftp.ics.uci.edu [128.195.1.1] in directory /pub/machine-learning-databases" or via web browser at http://www.ics.uci.edu/~mlearn/MLRepository.html

2. **UCI KDD Archive**

   The UC Irvine Knowledge Discovery in Databases (KDD) Archive at http://kdd.ics.uci.edu/ is an online repository of large datasets which encompasses a wide variety of data types, analysis tasks, and application areas. The primary role of this repository is to serve as a benchmark testbed to enable researchers in knowledge discovery and data mining to scale existing and future data analysis algorithms to very large and complex data sets. This archive is supported by the Information and Data Management Program at the National Science Foundation, and is intended to expand the current UCI Machine Learning Database Repository to datasets that are orders of magnitude larger and more complex.

3. **The neural-bench Benchmark collection**

   Accessible at http://www.boltz.cs.cmu.edu/ or via anonymous FTP at ftp://ftp.boltz.cs.cmu.edu/pub/neural-bench/. In case of problems or if you want to donate data, email contact is "neural-bench@cs.cmu.edu". The data sets in this repository include the 'nettalk' data, 'two spirals', protein structure prediction, vowel recognition, sonar signal classification, and a few others.

4. **Proben1**

   Proben1 is a collection of 12 learning problems consisting of real data. The datafiles all share a single simple common format. Along with the data comes a technical report describing a set of rules and conventions for performing and reporting benchmark tests and their results. Accessible via anonymous FTP on ftp.cs.cmu.edu [128.2.206.173] as /afs/cs/project/connect/bench/contrib/prechelt/proben1.tar.gz. and also on ftp.ira.uka.de as /pub/neuron/proben1.tar.gz. The file is about 1.8 MB and unpacks into about 20 MB.

5. **Delve: Data for Evaluating Learning in Valid Experiments**

   Delve is a standardised, copyrighted environment designed to evaluate the performance of learning methods. Delve makes it possible for users to compare their learning methods with other methods on many datasets. The Delve learning methods and evaluation procedures are well documented, such that meaningful comparisons can be made. The data collection includes not only isolated data sets, but "families" of data sets in which properties of the data, such as number of inputs and degree of nonlinearity or noise, are systematically varied. The Delve web page is at http://www.cs.toronto.edu/~delve/

6. **Bilkent University Function Approximation Repository**

   A repository of data sets collected mainly by searching resources on the web can be found at http://funapp.cs.bilkent.edu.tr/DataSets/ Most of the data sets are used for the experimental analysis of function approximation techniques and for training

and demonstration by machine learning and statistics community. The original sources of most data sets can be accessed via associated links. A compressed tar file containing all data sets is available.

7. **NIST special databases of the National Institute Of Standards And Technology:**

Several large databases, each delivered on a CD-ROM. Here is a quick list.
- NIST Binary Images of Printed Digits, Alphas, and Text
- NIST Structured Forms Reference Set of Binary Images
- NIST Binary Images of Handwritten Segmented Characters
- NIST 8-bit Gray Scale Images of Fingerprint Image Groups
- NIST Structured Forms Reference Set 2 of Binary Images
- NIST Test Data 1: Binary Images of Hand-Printed Segmented Characters
- NIST Machine-Print Database of Gray Scale and Binary Images
- NIST 8-Bit Gray Scale Images of Mated Fingerprint Card Pairs
- NIST Supplemental Fingerprint Card Data (SFCD) for NIST Special Database 9
- NIST Binary Image Databases of Census Miniforms (MFDB)
- NIST Mated Fingerprint Card Pairs 2 (MFCP 2)
- NIST Scoring Package Release 1.0
- NIST FORM-BASED HANDPRINT RECOGNITION SYSTEM

Here are example descriptions of two of these databases:

NIST special database 2: Structured Forms Reference Set (SFRS)

The NIST database of structured forms contains 5,590 full page images of simulated tax forms completed using machine print. THERE IS NO REAL TAX DATA IN THIS DATABASE. The structured forms used in this database are 12 different forms from the 1988, IRS 1040 Package X. These include Forms 1040, 2106, 2441, 4562, and 6251 together with Schedules A, B, C, D, E, F and SE. Eight of these forms contain two pages or form faces making a total of 20 form faces represented in the database. Each image is stored in bi-level black and white raster format. The images in this database appear to be real forms prepared by individuals but the images have been automatically derived and synthesized using a computer and contain no "real" tax data. The entry field values on the forms have been automatically generated by a computer in order to make the data available without the danger of distributing privileged tax information. In addition to the images the database includes 5,590 answer files, one for each image. Each answer file contains an ASCII representation of the data found in the entry fields on the corresponding image. Image format documentation and example software are also provided. The uncompressed database totals approximately 5.9 gigabytes of data.

NIST special database 3: Binary Images of Handwritten Segmented Characters (HWSC)

Contains 313,389 isolated character images segmented from the 2,100 full-page images distributed with "NIST Special Database 1". 223,125 digits, 44,951 upper-case, and 45,313 lower-case character images. Each character image has been centered in a separate 128 by 128 pixel region, error rate of the segmentation and assigned classification is less than 0.1%. The uncompressed database totals approximately 2.75 gigabytes of image data and includes image format documentation and example software.

The system requirements for all databases are a 5.25" CD-ROM drive with software to read ISO-9660 format. Contact: Darrin L. Dimmick; dld@magi.ncsl.nist.gov; (301)975-4147

The prices of the databases are between US$ 250 and 1895 If you wish to order a database, please contact: Standard Reference Data; National Institute of Standards and Technology; 221/A323; Gaithersburg, MD 20899; Phone: (301)975-2208; FAX: (301)926-0416

Samples of the data can be found by ftp on sequoyah.ncsl.nist.gov in directory /pub/data A more complete description of the available databases can be obtained from the same host as /pub/databases/catalog.txt

8. **CEDAR CD-ROM 1: Database of Handwritten Cities, States, ZIP Codes, Digits, and Alphabetic Characters**

The Center Of Excellence for Document Analysis and Recognition (CEDAR) State University of New York at Buffalo announces the availability of CEDAR CDROM 1: USPS Office of Advanced Technology The database contains handwritten words and ZIP Codes in high resolution grayscale (300 ppi 8-bit) as well as binary handwritten digits and alphabetic characters (300 ppi 1-bit). This database is intended to encourage research in off-line handwriting recognition by providing access to handwriting samples digitized from envelopes in a working post office.

```
        Specifications of the database include:
    +   300 ppi 8-bit grayscale handwritten words (cities,
        states, ZIP Codes)
```

```
            o    5632 city words
            o    4938 state words
            o    9454 ZIP Codes
    +    300 ppi binary handwritten characters and digits:
            o    27,837 mixed alphas  and  numerics  segmented
                 from address blocks
            o    21,179 digits segmented from ZIP Codes
    +    every image supplied with  a  manually  determined
         truth value
    +    extracted from live mail in a  working  U.S.  Post
         Office
    +    word images in the test  set  supplied  with  dic-
         tionaries  of  postal  words that simulate partial
         recognition of the corresponding ZIP Code.
    +    digit images included in test  set  that  simulate
         automatic ZIP Code segmentation.  Results on these
         data can be projected to overall ZIP Code recogni-
         tion performance.
    +    image format documentation and software included
```

System requirements are a 5.25" CD-ROM drive with software to read ISO-9660 format. For further information, see http://www.cedar.buffalo.edu/Databases/CDROM1/ or send email to Ajay Shekhawat at <ajay@cedar.Buffalo.EDU>

There is also a CEDAR CDROM-2, a database of machine-printed Japanese character images.

9. **AI-CD-ROM (see question "Other sources of information")**

10. **Time series**

Santa Fe Competition

Various datasets of time series (to be used for prediction learning problems) are available for anonymous ftp from ftp.santafe.edu in /pub/Time-Series". Data sets include:
- Fluctuations in a far-infrared laser
- Physiological data of patients with sleep apnea;
- High frequency currency exchange rate data;
- Intensity of a white dwarf star;
- J.S. Bachs final (unfinished) fugue from "Die Kunst der Fuge"

Some of the datasets were used in a prediction contest and are described in detail in the book "Time series prediction: Forecasting the future and understanding the past", edited by Weigend/Gershenfield, Proceedings Volume XV in the Santa Fe Institute Studies in the Sciences of Complexity series of Addison Wesley (1994).

M3 Competition

3003 time series from the M3 Competition can be found at http://forecasting.cwru.edu/Data/index.html

The numbers of series of various types are given in the following table:

Interval	Micro	Industry	Macro	Finance	Demog	Other	Total
Yearly	146	102	83	58	245	11	645
Quarterly	204	83	336	76	57	0	756
Monthly	474	334	312	145	111	52	1428
Other	4	0	0	29	0	141	174
Total	828	519	731	308	413	204	3003

Rob Hyndman's Time Series Data Library

A collection of over 500 time series on subjects including agriculture, chemistry, crime, demography, ecology, economics & finance, health, hydrology & meteorology, industry, physics, production, sales, simulated series, sport, transport & tourism, and tree-rings can be found at http://www-personal.buseco.monash.edu.au/~hyndman/TSDL/

11. **Financial data**

http://chart.yahoo.com/d?s=

http://www.chdwk.com/data/index.html

12. **USENIX Faces**

The USENIX faces archive is a public database, accessible by ftp, that can be of use to people working in the fields of human face recognition, classification and the like. It currently contains 5592 different faces (taken at USENIX conferences) and is updated twice each year. The images are mostly 96x128 greyscale frontal images and are stored in ascii files in a way that makes it easy to convert them to any usual graphic format (GIF, PCX, PBM etc.). Source code for viewers, filters, etc. is provided. Each image file takes approximately 25K.

For further information, see http://facesaver.usenix.org/

According to the archive administrator, Barbara L. Dijker (barb.dijker@labyrinth.com), there is no restriction to use them. However, the image files are stored in separate directories corresponding to the Internet site to which the person represented in the image belongs, with each directory containing a small number of images (two in the average). This makes it difficult to retrieve by ftp even a small part of the database, as you have to get each one individually.
A solution, as Barbara proposed me, would be to compress the whole set of images (in separate files of, say, 100 images) and maintain them as a specific archive for research on face processing, similar to the ones that already exist for fingerprints and others. The whole compressed database would take some 30 megabytes of disk space. I encourage anyone willing to host this database in his/her site, available for anonymous ftp, to contact her for details (unfortunately I don't have the resources to set up such a site).

Please consider that UUNET has graciously provided the ftp server for the FaceSaver archive and may discontinue that service if it becomes a burden. This means that people should not download more than maybe 10 faces at a time from uunet.

A last remark: each file represents a different person (except for isolated cases). This makes the database quite unsuitable for training neural networks, since for proper generalisation several instances of the same subject are required. However, it is still useful for use as testing set on a trained network.

13. **Linguistic Data Consortium**

The Linguistic Data Consortium (URL: http://www.ldc.upenn.edu/ldc/noframe.html) is an open consortium of universities, companies and government research laboratories. It creates, collects and distributes speech and text databases, lexicons, and other resources for research and development purposes. The University of Pennsylvania is the LDC's host institution. The LDC catalog includes pronunciation lexicons, varied lexicons, broadcast speech, microphone speech, mobile-radio speech, telephone speech, broadcast text, conversation text, newswire text, parallel text, and varied text, at widely varying fees.

```
Linguistic Data Consortium
University of Pennsylvania
3615 Market Street, Suite 200
Philadelphia, PA 19104-2608
Tel (215) 898-0464 Fax (215) 573-2175
Email: ldc@ldc.upenn.edu
```

14. **Otago Speech Corpus**

The Otago Speech Corpus contains speech samples in RIFF WAVE format that can be downloaded from http://divcom.otago.ac.nz/infosci/kel/software/RICBIS/hyspeech_main.html

15. **Astronomical Time Series**

Prepared by Paul L. Hertz (Naval Research Laboratory) & Eric D. Feigelson (Pennsyvania State University):
   - Detection of variability in photon counting observations 1 (QSO1525+337)
   - Detection of variability in photon counting observations 2 (H0323+022)
   - Detection of variability in photon counting observations 3 (SN1987A)
   - Detecting orbital and pulsational periodicities in stars 1 (binaries)
   - Detecting orbital and pulsational periodicities in stars 2 (variables)
   - Cross-correlation of two time series 1 (Sun)
   - Cross-correlation of two time series 2 (OJ287)
   - Periodicity in a gamma ray burster (GRB790305)
   - Solar cycles in sunspot numbers (Sun)
   - Deconvolution of sources in a scanning operation (HEAO A-1)
   - Fractal time variability in a seyfert galaxy (NGC5506)
   - Quasi-periodic oscillations in X-ray binaries (GX5-1)
   - Deterministic chaos in an X-ray pulsar? (Her X-1)
URL: http://xweb.nrl.navy.mil/www_hertz/timeseries/timeseries.html

16. **Miscellaneous Images**

The USC-SIPI Image Database: http://sipi.usc.edu/services/database/Database.html

CityU Image Processing Lab: http://www.image.cityu.edu.hk/images/database.html

Center for Image Processing Research: http://cipr.rpi.edu/

Computer Vision Test Images: http://www.cs.cmu.edu:80/afs/cs/project/cil/ftp/html/v-images.html

Lenna 97: A Complete Story of Lenna: http://www.image.cityu.edu.hk/images/lenna/Lenna97.html

17. **StatLib**

The StatLib repository at http://lib.stat.cmu.edu/ at Carnegie Mellon University has a large collection of data sets, many of which can be used with NNs.

Next part is part 5 (of 7). Previous part is part 3.