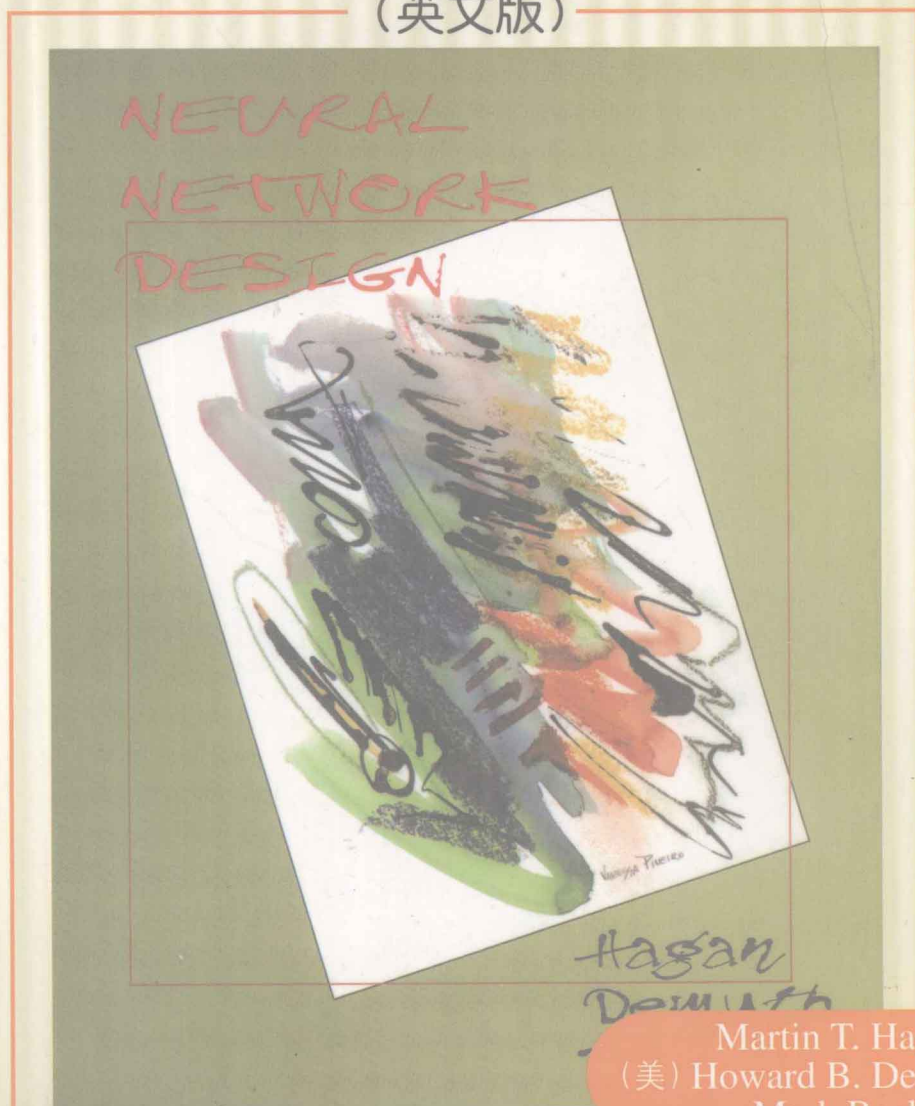


经 典 原 版 书 库

神经网络设计

(英文版)



Martin T. Hagan
(美) Howard B. Demuth 著
Mark Beale



机械工业出版社
China Machine Press



中信出版社
CITIC PUBLISHING HOUSE

THOMSON
T.M.

经典原版书库

神经网络设计

(英文版)

Neural Network Design

Martin T. Hagan
(美) Howard B. Demuth 著
Mark Beale



机械工业出版社
China Machine Press



中信出版社
CITIC PUBLISHING HOUSE

Martin T. Hagan, Howard B. Demuth: Neural Network Design

Original copyright © 1996 by PWS Publishing Company. All rights reserved.

First published by PWS Publishing Company, a division of Thomson Learning, United States of America.

Reprinted for People's Republic of China by Thomson Asia Pte Ltd and China Machine Press and CITIC Publishing House under the authorization of Thomson Learning. No part of this book may be reproduced in any form without the the prior written permission of Thomson Learning and China Machine Press.

本书影印版由美国汤姆森学习出版集团授权机械工业出版社和中信出版社出版。未经出版者书面许可，不得以任何方式复制或抄袭本书内容。

版权所有，侵权必究。

本书版权登记号：图字：01-2001-5321

图书在版编目（CIP）数据

神经网络设计/（美）黑根（Hagan, M. T.）等著. - 北京：机械工业出版社，2002.8
（经典原版书库）

书名原文：Neural Network Design

ISBN 7-111-10841-8

I. 神… II. 黑… III. 人工神经网络 - 系统设计 - 英文 IV. TP183

中国版本图书馆CIP数据核字（2002）第063186号

机械工业出版社（北京市西城区百万庄大街22号 邮政编码 100037）

责任编辑：华章

北京忠信诚印刷厂印刷·新华书店北京发行所发行

2002年8月第1版第1次印刷

787mm × 1092mm 1/16 · 46印张

印数：0 001-3 000册

定价：69.00元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换

Preface

This book gives an introduction to basic neural network architectures and learning rules. Emphasis is placed on the mathematical analysis of these networks, on methods of training them and on their application to practical engineering problems in such areas as pattern recognition, signal processing and control systems.

Every effort has been made to present material in a clear and consistent manner so that it can be read and applied with ease. We have included many solved problems to illustrate each topic of discussion.

Since this is a book on the design of neural networks, our choice of topics was guided by two principles. First, we wanted to present the most useful and practical neural network architectures, learning rules and training techniques. Second, we wanted the book to be complete in itself and to flow easily from one chapter to the next. For this reason, various introductory materials and chapters on applied mathematics are included just before they are needed for a particular subject. In summary, we have chosen some topics because of their practical importance in the application of neural networks, and other topics because of their importance in explaining how neural networks operate.

We have omitted many topics that might have been included. We have not, for instance, made this book a catalog or compendium of all known neural network architectures and learning rules, but have instead concentrated on the fundamental concepts. Second, we have not discussed neural network implementation technologies, such as VLSI, optical devices and parallel computers. Finally, we do not present the biological and psychological foundations of neural networks in any depth. These are all important topics, but we hope that we have done the reader a service by focusing on those topics that we consider to be most useful in the design of neural networks and by treating those topics in some depth.

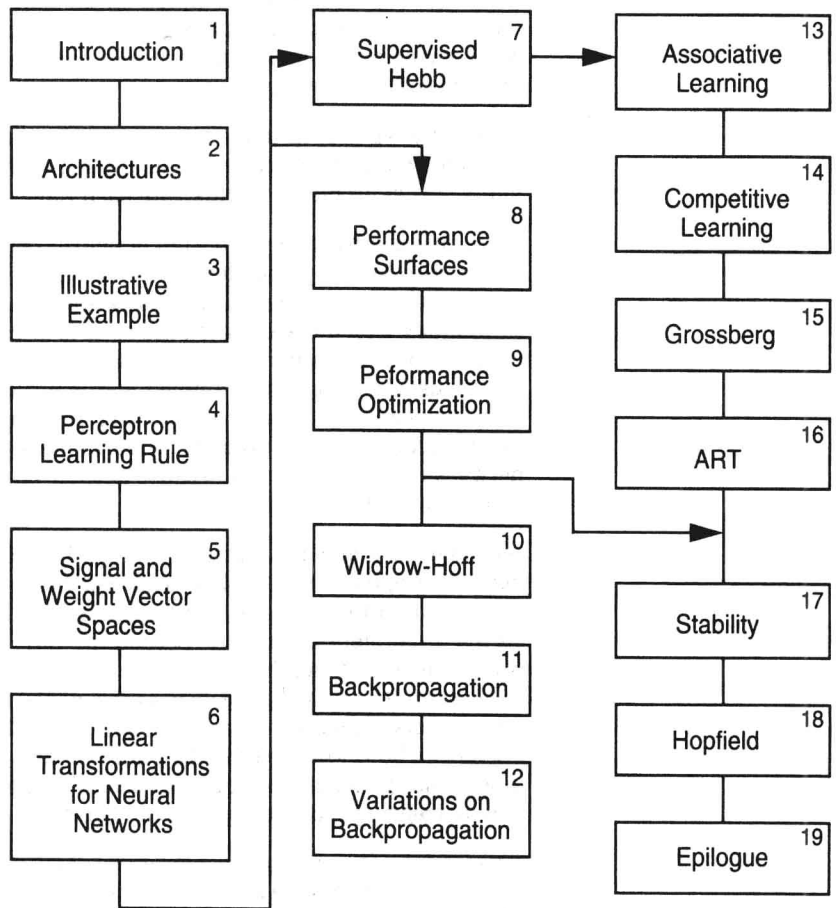
This book has been organized for a one-semester introductory course in neural networks at the senior or first-year graduate level. (It is also suitable for short courses, self-study and reference.) The reader is expected to have some background in linear algebra, probability and differential equations.

Each chapter of the book is divided into the following sections: Objectives, Theory and Examples, Summary of Results, Solved Problems, Epilogue,



Further Reading and Exercises. The *Theory and Examples* section comprises the main body of each chapter. It includes the development of fundamental ideas as well as worked examples (indicated by the icon shown here in the left margin). The *Summary of Results* section provides a convenient listing of important equations and concepts and facilitates the use of the book as an industrial reference. About a third of each chapter is devoted to the *Solved Problems* section, which provides detailed examples for all key concepts.

The following figure illustrates the dependencies among the chapters.



Chapters 1 through 6 cover basic concepts that are required for all of the remaining chapters. Chapter 1 is an introduction to the text, with a brief historical background and some basic biology. Chapter 2 describes the basic neural network architectures. The notation that is introduced in this chapter is used throughout the book. In Chapter 3 we present a simple pattern recognition problem and show how it can be solved using three different types of neural networks. These three networks are representative of the types of networks that are presented in the remainder of the text. In addition, the pattern recognition problem presented here provides a common thread of experience throughout the book.

Much of the focus of this book will be on methods for training neural networks to perform various tasks. In Chapter 4 we introduce learning algorithms and present the first practical algorithm: the perceptron learning rule. The perceptron network has fundamental limitations, but it is important for historical reasons and is also a useful tool for introducing key concepts that will be applied to more powerful networks in later chapters.

One of the main objectives of this book is to explain how neural networks operate. For this reason we will weave together neural network topics with important introductory material. For example, linear algebra, which is the core of the mathematics required for understanding neural networks, is reviewed in Chapters 5 and 6. The concepts discussed in these chapters will be used extensively throughout the remainder of the book.

Chapters 7 and 13–16 describe networks and learning rules that are heavily inspired by biology and psychology. They fall into two categories: associative networks and competitive networks. Chapters 7 and 13 introduce basic concepts, while Chapters 14–16 describe more advanced networks.

Chapters 8–12 develop a class of learning called performance learning, in which a network is trained to optimize its performance. Chapters 8 and 9 introduce the basic concepts of performance learning. Chapters 10–12 apply these concepts to feedforward neural networks of increasing power and complexity.

Chapters 17 and 18 discuss recurrent networks. These networks, which have feedback connections, are dynamical systems. Chapter 17 investigates the stability of these systems. Chapter 18 presents the Hopfield network, which has been one of the most influential recurrent networks.

In Chapter 19 we summarize the networks presented in this book and discuss their relationships to other networks that we do not cover. We also point the reader to other sources for further study. If you want to know “Where do I go from here?” look to Chapter 19.

Software

MATLAB is not essential for using this book. The computer exercises can be performed with any available programming language, and the *Neural Network Design Demonstrations*, while helpful, are not critical to understanding the material covered in this book.

However, we have made use of the MATLAB software package to supplement the textbook. This software is widely available and, because of its matrix/vector notation and graphics, is a convenient environment in which to experiment with neural networks. We use MATLAB in two different ways. First, we have included a number of exercises for the reader to perform in MATLAB. Many of the important features of neural networks become apparent only for large-scale problems, which are computationally intensive and not feasible for hand calculations. With MATLAB, neural network algorithms can be quickly implemented, and large-scale problems can be tested conveniently. These MATLAB exercises are identified by the icon shown here to the left. (If MATLAB is not available, any other programming language can be used to perform the exercises.)



The second way in which we use MATLAB is through the *Neural Network Design Demonstrations*, which are on a disk included with this book. These interactive demonstrations illustrate important concepts in each chapter. After the software has been loaded into the MATLAB directory on your computer, it can be invoked by typing `nnd` at the MATLAB prompt. All demonstrations are easily accessible from a master menu. The icon shown here to the left identifies references to these demonstrations in the text.



The demonstrations require MATLAB version 4.0 or later, or the student edition of MATLAB version 4.0. In addition, a few of the demonstrations require The MathWorks' *Neural Network Toolbox* version 1.0 or later. See Appendix C for specific information on using the demonstration software.

As an aid to instructors who are using this text, we have prepared a companion set of overheads. Transparency masters (in Microsoft Powerpoint format) for each chapter are available on the web at: www.pws.com/pwsftp.html. A solutions manual is also available. Contact PWS Publishing at (800) 347-7707 for more information.

Acknowledgments

We are deeply indebted to the reviewers who have given freely of their time to read all or parts of the drafts of this book and to test various versions of the software. In particular we are most grateful to Professor John Andreae, University of Canterbury; Dan Foresee, AT&T; Dr. Carl Latino, Oklahoma State University; Jack Hagan, MCI; Dr. Gerry Andeen, SRI; and Joan Miller and Margie Jenks, University of Idaho. We also had constructive inputs from our graduate students in ECEN 5713 at Oklahoma State University and ENEL 621 at the University of Canterbury who read early drafts, tested the software and provided helpful suggestions for improving the book. We are also grateful to the anonymous reviewers who provided several useful recommendations.

We wish to thank Dr. Peter Gough for inviting us to join the staff in the Electrical and Electronic Engineering Department at the University of Canterbury, Christchurch, New Zealand. Thanks also to Mike Surety for his computer help and to the departmental staff for their assistance. A sabbatical from Oklahoma State University and a year's leave from the University of Idaho gave us the time to write this book. Thanks to Texas Instruments, and in particular Bill Harland, for their support of our neural network research. Thanks to The Mathworks for permission to use material from the *Neural Network Toolbox*.

We are grateful to Joan Pilgram for her encouragement and business advice, and to Mrs. Bernice Hewitt, Christchurch, for her good spirit and hospitality.

Finally, we wish to express our appreciation to the staff at PWS Publishing Company, especially Bill Barter, Pam Rockwell, Amy Mayfield, Ken Morton and Nathan Wilbur. Thanks to Vanessa Pineiro for the lovely cover art.

Contents

Preface

1

Introduction

Objectives	1-1
History	1-2
Applications	1-5
Biological Inspiration	1-8
Further Reading	1-10

2

Neuron Model and Network Architectures



Objectives	2-1
Theory and Examples	2-2
Notation	2-2
Neuron Model	2-2
Single-Input Neuron	2-2
Transfer Functions	2-3
Multiple-Input Neuron	2-7
Network Architectures	2-9
A Layer of Neurons	2-9
Multiple Layers of Neurons	2-10
Recurrent Networks	2-13
Summary of Results	2-16
Solved Problems	2-20
Epilogue	2-22
Exercises	2-23

3**An Illustrative Example**

Objectives	3-1
Theory and Examples	3-2
Problem Statement	3-2
Perceptron	3-3
Two-Input Case	3-4
Pattern Recognition Example	3-5
Hamming Network	3-8
Feedforward Layer	3-8
Recurrent Layer	3-9
Hopfield Network	3-12
Epilogue	3-15
Exercise	3-16

4**Perceptron Learning Rule**

Objectives	4-1
Theory and Examples	4-2
Learning Rules	4-2
Perceptron Architecture	4-3
Single-Neuron Perceptron	4-5
Multiple-Neuron Perceptron	4-8
Perceptron Learning Rule	4-8
Test Problem	4-9
Constructing Learning Rules	4-10
Unified Learning Rule	4-12
Training Multiple-Neuron Perceptrons	4-13
Proof of Convergence	4-15
Notation	4-15
Proof	4-16
Limitations	4-18
Summary of Results	4-20
Solved Problems	4-21
Epilogue	4-33
Further Reading	4-34
Exercises	4-36

5**Signal and Weight Vector Spaces**

Objectives	5-1
Theory and Examples	5-2
Linear Vector Spaces	5-2
Linear Independence	5-4
Spanning a Space	5-5
Inner Product	5-6
Norm	5-7
Orthogonality	5-7
Gram-Schmidt Orthogonalization	5-8
Vector Expansions	5-9
Reciprocal Basis Vectors	5-10
Summary of Results	5-14
Solved Problems	5-17
Epilogue	5-26
Further Reading	5-27
Exercises	5-28

6**Linear Transformations for Neural Networks**

Objectives	6-1
Theory and Examples	6-2
Linear Transformations	6-2
Matrix Representations	6-3
Change of Basis	6-6
Eigenvalues and Eigenvectors	6-10
Diagonalization	6-13
Summary of Results	6-15
Solved Problems	6-17
Epilogue	6-28
Further Reading	6-29
Exercises	6-30

7**Supervised Hebbian Learning**

Objectives	7-1
Theory and Examples	7-2
Linear Associator	7-3
The Hebb Rule	7-4
Performance Analysis	7-5
Pseudoinverse Rule	7-7
Application	7-10
Variations of Hebbian Learning	7-12
Summary of Results	7-14
Solved Problems	7-16
Epilogue	7-29
Further Reading	7-30
Exercises	7-31

8**Performance Surfaces and Optimum Points**

Objectives	8-1
Theory and Examples	8-2
Taylor Series	8-2
Vector Case	8-4
Directional Derivatives	8-5
Minima	8-7
Necessary Conditions for Optimality	8-9
First-Order Conditions	8-10
Second-Order Conditions	8-11
Quadratic Functions	8-12
Eigensystem of the Hessian	8-13
Summary of Results	8-20
Solved Problems	8-22
Epilogue	8-34
Further Reading	8-35
Exercises	8-36

9**Performance Optimization**

Objectives	9-1
Theory and Examples	9-2
Steepest Descent	9-2
Stable Learning Rates	9-6
Minimizing Along a Line	9-8
Newton's Method	9-10
Conjugate Gradient	9-15
Summary of Results	9-21
Solved Problems	9-23
Epilogue	9-37
Further Reading	9-38
Exercises	9-39

10**Widrow-Hoff Learning**

Objectives	10-1
Theory and Examples	10-2
ADALINE Network	10-2
Single ADALINE	10-3
Mean Square Error	10-4
LMS Algorithm	10-7
Analysis of Convergence	10-9
Adaptive Filtering	10-13
Adaptive Noise Cancellation	10-15
Echo Cancellation	10-21
Summary of Results	10-22
Solved Problems	10-24
Epilogue	10-40
Further Reading	10-41
Exercises	10-42

11

Backpropagation



Objectives	11-1
Theory and Examples	11-2
Multilayer Perceptrons	11-2
Pattern Classification	11-3
Function Approximation	11-4
The Backpropagation Algorithm	11-7
Performance Index	11-8
Chain Rule	11-9
Backpropagating the Sensitivities	11-11
Summary	11-13
Example	11-14
Using Backpropagation	11-17
Choice of Network Architecture	11-17
Convergence	11-19
Generalization	11-21
Summary of Results	11-24
Solved Problems	11-26
Epilogue	11-40
Further Reading	11-41
Exercises	11-43

12

Variations on Backpropagation



Objectives	12-1
Theory and Examples	12-2
Drawbacks of Backpropagation	12-3
Performance Surface Example	12-3
Convergence Example	12-7
Heuristic Modifications of Backpropagation	12-9
Momentum	12-9
Variable Learning Rate	12-12
Numerical Optimization Techniques	12-14
Conjugate Gradient	12-14
Levenberg-Marquardt Algorithm	12-19
Summary of Results	12-28
Solved Problems	12-32
Epilogue	12-46
Further Reading	12-47
Exercises	12-50

13

Associative Learning



Objectives	13-1
Theory and Examples	13-2
Simple Associative Network	13-3
Unsupervised Hebb Rule	13-5
Hebb Rule with Decay	13-7
Simple Recognition Network	13-9
Instar Rule	13-11
Kohonen Rule	13-15
Simple Recall Network	13-16
Outstar Rule	13-17
Summary of Results	13-21
Solved Problems	13-23
Epilogue	13-34
Further Reading	13-35
Exercises	13-37

14

Competitive Networks



Objectives	14-1
Theory and Examples	14-2
Hamming Network	14-3
Layer 1	14-3
Layer 2	14-4
Competitive Layer	14-5
Competitive Learning	14-7
Problems with Competitive Layers	14-9
Competitive Layers in Biology	14-10
Self-Organizing Feature Maps	14-12
Improving Feature Maps	14-15
Learning Vector Quantization	14-16
LVQ Learning	14-18
Improving LVQ Networks (LVQ2)	14-21
Summary of Results	14-22
Solved Problems	14-24
Epilogue	14-37
Further Reading	14-38
Exercises	14-39

Grossberg Network

Objectives	15-1
Theory and Examples	15-2
Biological Motivation: Vision	15-3
Illusions	15-4
Vision Normalization	15-8
Basic Nonlinear Model	15-9
Two-Layer Competitive Network	15-12
Layer 1	15-13
Layer 2	15-17
Choice of Transfer Function	15-20
Learning Law	15-22
Relation to Kohonen Law	15-24
Summary of Results	15-26
Solved Problems	15-30
Epilogue	15-42
Further Reading	15-43
Exercises	15-45

Adaptive Resonance Theory

Objectives	16-1
Theory and Examples	16-2
Overview of Adaptive Resonance	16-2
Layer 1	16-4
Steady State Analysis	16-6
Layer 2	16-10
Orienting Subsystem	16-13
Learning Law: L1-L2	16-17
Subset/Superset Dilemma	16-17
Learning Law	16-18
Learning Law: L2-L1	16-20
ART1 Algorithm Summary	16-21
Initialization	16-21
Algorithm	16-21
Other ART Architectures	16-23
Summary of Results	16-25
Solved Problems	16-30
Epilogue	16-45
Further Reading	16-46
Exercises	16-48

17

Stability



Objectives	17-1
Theory and Examples	17-2
Recurrent Networks	17-2
Stability Concepts	17-3
Definitions	17-4
Lyapunov Stability Theorem	17-5
Pendulum Example	17-6
LaSalle's Invariance Theorem	17-12
Definitions	17-12
Theorem	17-13
Example	17-14
Comments	17-18
Summary of Results	17-19
Solved Problems	17-21
Epilogue	17-28
Further Reading	17-29
Exercises	17-30

18

Hopfield Network



Objectives	18-1
Theory and Examples	18-2
Hopfield Model	18-3
Lyapunov Function	18-5
Invariant Sets	18-7
Example	18-7
Hopfield Attractors	18-11
Effect of Gain	18-12
Hopfield Design	18-16
Content-Addressable Memory	18-16
Hebb Rule	18-18
Lyapunov Surface	18-22
Summary of Results	18-24
Solved Problems	18-26
Epilogue	18-36
Further Reading	18-37
Exercises	18-40