

国际知名大学原版教材

—— 信息技术学科与电气工程学科系列

12

NEURAL NETWORKS
A Comprehensive Foundation
Second Edition

神经网络的综合基础

第 2 版

Simon Haykin



清华大学出版社

<http://www.tup.tsinghua.edu.cn>

Prentice
Hall

培生教育出版集团

<http://www.pearsoned.com>

NEURAL NETWORKS

A Comprehensive Foundation

Second Edition

Simon Haykin

*McMaster University
Hamilton, Ontario, Canada*

Tsinghua University Press

Prentice Hall

(京) 新登字 158 号

Neural Networks: A Comprehensive Foundation, Second Edition

Copyright © 1999 by Prentice-Hall, Inc.

Original English language Edition Published by Prentice-Hall, Inc.

All Rights Reserved.

For sales in Mainland China only.

本书影印版由培生教育出版集团授权清华大学出版社在中国境内（不包括香港、澳门特别行政区和台湾地区）独家出版、发行。

未经出版者书面许可，不得以任何方式复制或抄袭本书的任何部分。

本书封面贴有培生教育出版集团防伪标签，无标签者不得销售。

北京市版权局著作权合同登记号：图字 01-2001-3791

书 名：神经网络的综合基础（第 2 版）

作 者：Simon Haykin

出版者：清华大学出版社(北京清华大学学研大厦，邮编 100084)

[http:// www.tup.tsinghua.edu.cn](http://www.tup.tsinghua.edu.cn)

印刷者：北京市四季青印刷厂

发行者：新华书店总店北京发行所

开 本：787×960 1/16 印张：54.75

版 次：2001 年 10 月第 1 版 2001 年 10 月第 1 次印刷

书 号：ISBN 7-302-04936-X/TP · 2778

印 数：0001~3000

定 价：78.00 元

国际知名大学原版教材

——信息技术学科与电气工程学科系列

出版说明

郑大钟

清华大学信息技术科学与技术学院

当前,在我国的高等学校中,教学内容和课程体系的改革已经成为教学改革中的一个非常突出的问题,而为数不少的课程教材中普遍存在“课程体系老化,内容落伍时代,本研层次不清”的现象又是其中的急需改变的一个重要方面。同时,随着科教兴国方针的贯彻落实,要求我们进一步转变观念扩大视野,使教学过程适应以信息技术为先导的技术革命和我国社会主义市场经济的需要,加快教学过程的国际化进程。在这方面,系统地研究和借鉴国外知名大学的相关教材,将会对推进我们的课程改革和推进我国大学教学的国际化进程,乃至对我们一些重点大学建设国际一流大学的努力,都将具有重要的借鉴推动作用。正是基于这种背景,我们决定在国内推出信息技术学科和电气工程学科国外知名大学原版系列教材。

本系列教材的组编将遵循如下的几点基本原则。(1)书目的范围限于信息技术学科和电气工程学科所属专业的技术基础课和主要的专业课。(2)教材的范围选自于具有较大影响且为国外知名大学所采用的教材。(3)教材属于在近5年内所出版的新书或新版书。(4)教材适合于作为我国大学相应课程的教材或主要教学参考书。(5)每本列选的教材都须经过国内相应领域的资深专家审看和推荐。(6)教材的形式直接以英文原版形式印刷出版。

本系列教材将按分期分批的方式组织出版。为了便于使用本系列教材的相关教师和学生从学科和教学的角度对其在体系和内容上的特点和特色有所了解,在每本教材中都附有我们所约请的相关领域资深教授撰写的影印版序言。此外,出于多样化的考虑,对于某些基本类型的课程,我们还同时列选了多于一本的不同体系、不同风格 and 不同层次的教材,以供不同要求和不同学时的同类课程的选用。

本系列教材的读者对象为信息技术学科和电气工程学科所属各专业的本科生,同时兼顾其他工程学科专业的本科生或研究生。本系列教材,既可采用作为相应课程的教材或教学参考书,也可提供作为工作于各个技术领域的工程师和技术人员的自学读物。

组编这套国外知名大学原版系列教材是一个尝试。不管是书目确定的合理性,教材选择的恰当性,还是评论看法的确切性,都有待于通过使用和实践来检验。感谢使用本系列教材的广大教师和学生的支持。期望广大读者提出意见和建议。

46 5/04

Neural Networks: A Comprehensive Foundation,

Second Edition

影印版序

介绍新兴学科内容的书或教科书不少，但是，能在国际学术界得到广泛赞誉的书却不多。加拿大 McMaster 大学 S. Haykin 教授的《Neural Networks: A Comprehensive Foundation》就是这样一本珍贵的世界名著。一方面，它从应用和工程的角度，对神经计算的基础进行了广泛而深刻的论述；同时，又包容了该学科最新的研究成果。该书于 1994 年首次出版后不久，就受到学术界广泛重视，引用它的文献迅速增加，越来越多的高校以它为教材。该领域不少世界顶尖科学家参与了该书的校阅。

作者 S. Haykin 是一位知识渊博的科学家，一位能在物理学、数学、信号处理和通信等知识领域自如驾驭的舵手。他著作或编辑的书有 50 多本，其中《Adaptive Filter Theory》，《Communication Systems》，《Signal and System》已闻名于世，他发表了 500 多篇论文，拥有 3 项专利。作为一位多产的著作家，前期在自适应滤波与通信理论方面出版了优秀的教科书，随后又成功地把多种快速发展的技术的重要思想融会到这一本简明的、综合性教材与文献之中，使它成为科技工作者和研究生进入这一领域的理想导论。该书详细分析了最知名的人工神经网络（ANN）模型，对它们进行了富有成果的比较，尤其注意从系统论和统计学的角度分析问题。

本书第 1 章是一个清晰的引言，第 2 章提出学习过程及其统计特性的分类方法。这引导性的两章涉及众多的 ANN 基本概念，应引起读者的足够重视。第 3 章到第 7 章论述有指导的学习机；第 8 至 12 章是关于无指导的学习机；第 13 至 15 章讨论非线性动力学系统，包括前馈网络及回复式网络。该书的尾声简要地叙述了神经网络在构建用于模式识别、控制和信号处理的智能机器中所扮演的角色。

作者在该书的序言中对书的内容安排作了简明介绍，并指出，书中多数章节可以选读，作者还在附加的清单中为如何取舍提出了参考意见，提供了网址以供查找更多的有关信息。

近些年来，这一事实已被公认：神经计算不仅是数据驱动问题求解的有力工具，而且作为计算智能的核心内容之一，它也是不确定性、非线性或复杂性问题求解的有力工具。因此，一本系统的、综合最新研究成果的、指导性的书，无论对我国的研究生还是对科技工作者，都是急需的。该书从应用和工程的角度出发，但又注重算法分析而不是应用，从而大大提高了它作为教科书的价值。

笔者于 2000 年 1 月与 S. Haykin 教授商量翻译并在我国出版本书的问题，他

欣然同意。今年，清华大学出版社与 Prentice-Hall 达成协议，考虑本书主要读者的特点，决定由清华大学出版社原文再版，面向我国相关学科的研究生、科技工作者及其他对此领域感兴趣的人士。在此，谨代表广大读者对他们的诚意与协同合作表示衷心感谢。

冯天瑾
青岛海洋大学信息工程中心教授
2001 年 7 月

Neural Networks: A Comprehensive Foundation

Preface

It is a distinct honour for me to write this Preface for the Chinese edition of my book *Neural Networks: A Comprehensive Foundation* (2nd edition, Prentice-Hall, 1999), which is photocopy published by Tsinghua University Press in China, under agreement between Prentice-Hall in the United States and Tsinghua University Press. The publication of this edition is all the more pleasurable for me as the book will be read by a very large body of people in the big country of China.

Neural Networks represent a technology that is rooted in many disciplines: neuro-sciences, mathematics, statistics, physics, computer science, and engineering. As such, neural networks find applications in such diverse fields as system modeling, time series analysis, pattern recognition, signal processing, and control. They do so by virtue of their ability to learn (in on-line or off-line ways) from input data, representative of an environment of interest, with or without a teacher. Neural networks, motivated by how the human brain works, will play an increasingly important role in the twenty first century, particularly when we are confronted with difficult problems that are characterized by non-linearity, non-stationarity, and unknown statistics.

As the title of the book indicates, the primary purpose of this book is to provide a comprehensive introduction to neural networks from an engineering perspective, recognizing the multi-disciplinary nature of this fascinating subject. The material presented in the fifteen chapters of the book is supplied with numerous examples, computer-oriented experiments, end-of-chapter problems, and an extensive bibliography. This book is well organized and clearly written, trying to clarify the most important concepts and theory in the neural network (connectionist) literature. The original English edition of the book has been well received throughout the world; hopefully, the Chinese edition of the book will be just as well received in China.

In a related way, I would like to bring to the attention of potential readers my other book, *Adaptive Filter Theory* (3rd edition, Prentice-Hall, 1996); the fourth edition of this highly popular book will have been published before the end of 2001. The two books, *Neural Networks* and *Adaptive Filter Theory*, complement each other; between them, they cover all the important principles of Adaptive and Learning

Systems.

Finally, I would like to express my gratitude to those who worked in the production of the photocopied version of my book in China. Special thanks go to Prof. Tian-Jin Feng and my current Chinese Ph. D. student, Zhe Chen, who were the first to make this proposal and finally make it happen.

Simon Haykin
McMaster University
Hamilton, Ontario, Canada

To the countless researchers in neural networks
for their original contributions,
the many reviewers for their critical inputs,
my many graduate students for their keen interest,
and
my wife, Nancy, for her patience and tolerance.

Preface

Neural Networks, or artificial neural networks to be more precise, represent a technology that is rooted in many disciplines: neurosciences, mathematics, statistics, physics, computer science, and engineering. Neural networks find applications in such diverse fields as modeling, time series analysis, pattern recognition, signal processing, and control by virtue of an important property: the ability to *learn* from input data with or without a teacher.

This book provides a comprehensive foundation of neural networks, recognizing the multidisciplinary nature of the subject. The material presented in the book is supported with examples, computer-oriented experiments, end-of-chapter problems, and a bibliography.

The book consists of four parts, organized as follows:

1. *Introductory material*, consisting of Chapters 1 and 2. Chapter 1 describes, largely in qualitative terms, what neural networks are, their properties, compositions, and how they relate to artificial intelligence. This chapter ends with some historical notes. Chapter 2 provides an overview of the many facets of the learning process and its statistical properties. This chapter introduces an important concept: the *Vapnik–Chervonenkis (VC) dimension* used as a measure for the capacity of a family of classification functions realized by a learning machine.
2. *Learning machines with a teacher*, consisting of Chapters 3 through 7. Chapter 3 studies the simplest class of neural networks in this part: networks involving one or more output neurons but no hidden ones. The least-mean-square (LMS) algorithm (highly popular in the design of linear adaptive filters) and the perceptron-convergence theorem are described in this chapter. Chapter 4 presents an exhaustive treatment of *multilayer perceptrons* trained with the *back-propagation algorithm*. This algorithm (representing a generalization of the LMS algorithm) has emerged as the workhorse of neural networks. Chapter 5 presents detailed mathematical treatment of another class of layered neural networks: *radial-basis function networks*, whose composition involves a single layer of basis functions. This chapter emphasizes the role of regularization theory in the design of RBF

networks. Chapter 6 describes a relatively new class of learning machines known as *support vector machines*, whose theory builds on the material presented in Chapter 2 on statistical learning theory. The second part of the book finishes in Chapter 7 with a discussion of *committee machines*, whose composition involves several learners as components. In this chapter we describe *ensemble averaging*, *boosting*, and *hierarchical mixture of experts* as three different methods of building a committee machine.

3. *Learning machines without a teacher*, consisting of Chapters 8 through 12. Chapter 8 applies *Hebbian learning* to *principal components analysis*. Chapter 9 applies another form of self-organized learning, namely *competitive learning*, to the construction of computational maps known as *self-organizing maps*. These two chapters distinguish themselves by emphasizing learning rules that are rooted in neurobiology. Chapter 10 looks to *information theory* for the formulation of unsupervised learning algorithms, and emphasizes their applications to *modeling*, *image processing*, and *independent components analysis*. Chapter 11 describes self-supervised learning machines rooted in *statistical mechanics*, a subject that is closely allied to information theory. Chapter 12, the last chapter in the third part of the book, introduces *dynamic programming* and its relationship to *reinforcement learning*.
4. *Nonlinear dynamical systems*, consisting of Chapters 13 through 15. Chapter 13 describes a class of dynamical systems consisting of short-term memory and layered feedforward network structures. Chapter 14 emphasizes the issue of stability that arises in nonlinear dynamical systems involving the use of *feedback*. Examples of *associative memory* are discussed in this chapter. Chapter 15 describes another class of nonlinear dynamical systems, namely *recurrent networks*, that rely on the use of feedback for the purpose of input–output mapping.

The book concludes with an epilogue that briefly describes the role of neural networks in the construction of *intelligent machines* for pattern recognition, control, and signal processing.

The organization of the book offers a great deal of flexibility for use in graduate courses on neural networks. The final selection of topics can only be determined by the interests of the instructors using the book. To help in this selection process, a study guide is included in the accompanying manual.

There are a total of 15 computer-oriented experiments distributed throughout the book. Thirteen of these experiments use MATLAB. The files for the MATLAB experiments can be directly downloaded from

<ftp://ftp.mathworks.com/pub/books/haykin>

or alternatively

<http://www.mathworks.com/books/>

In this second case, the user will have to click on “Neural/Fuzzy” and then on the title of the book. The latter approach provides a nicer interface.

Each chapter ends with a set of problems. Many of the problems are of a challenging nature, designed not only to test the user of the book for how well the material

covered in the book has been understood, but also to extend that material. Solutions to all of the problems are described in an accompanying manual. Copies of this manual are only available to instructors who adopt the book, which can be obtained by writing to the publisher of the book, Prentice Hall.

The book should appeal to engineers, computer scientists, and physicists. It is hoped that researchers in other disciplines such as psychology and neurosciences will also find the book useful.

Simon Haykin
Hamilton, Ontario
February, 1998.

Acknowledgments

I am deeply indebted to the many reviewers who have given freely of their time to read through the book, in part or in full. In particular, I would like to express my deep gratitude to Dr. Kenneth Rose, University of California at Santa Barbara, for his many constructive inputs and invaluable help.

I am grateful to Dr. S. Amari, RIKEN, Japan; Dr. Sue Becker, McMaster University; Dr. Ron Racine, McMaster University; Dr. Sean Holden, University College, London; Dr. Michael Turmon, JPL, Pasadena; Dr. Babak Hassibi, Stanford University; Dr. Paul Yee, formerly of McMaster University; Dr. Edgar Osuna, MIT; Dr. Bernard Schölkopf, Max Planck Institute, Germany; Dr. Michael Jordan, MIT; Dr. Radford Neal, University of Toronto; Dr. Zoubin Gharhamani, University of Toronto; Dr. Marc Van Hulle, Katholieke Universiteit Leuven, Belgium; Dr. John Tsitsiklis, MIT; Dr. Jose Principe, University of Florida, Gainesville; Mr. Gint Puskorius, Ford Research Laboratory, Dearborn, Mich.; Dr. Lee Feldkamp, Ford Research Laboratory, Dearborn, Mich.; Dr. Lee Giles, NEC Research Institute, Princeton, NJ; Dr. Mikel Forcada, Universitat d'Alacant, Spain; Dr. Eric Wan, Oregon Graduate Institute of Science and Technology; Dr. Yann LeCun, AT&T Research, NJ; Dr. Jean-Francois Cardoso, Ecole Nationale, Paris; Dr. Anthony Bell, formerly of Salk Institute, San Diego; and Dr. Stefan Kremer, University of Guelph. They all helped me immeasurably in improving the presentation of material in different parts of the book.

I also wish to thank Dr. Ralph Linsker, IBM, Watson Research Center; Dr. Yaser Abu-Mostafa, Cal Tech.; Dr. Stuart Geman, Brown University; Dr. Alan Gelford, University of Connecticut; Dr. Yoav Freund, AT&T Research; Dr. Bart Kosko, University of Southern California; Dr. Narish Sinha, McMaster University; Dr. Grace Wahba, University of Wisconsin; Dr. Kostas Diamantaras, Aristotelian University of Thessaloniki, Greece; Dr. Robert Jacobs, University of Rochester; Dr. Peter Dayan, MIT; Dr. Dimitris Bertsekas, MIT; Dr. Andrew Barto, University of Massachusetts; Dr. Don Hush, University of New Mexico; Dr. Yoshua Bengio, University of Montreal; Dr. Andrew Cichoki, RIKEN, Japan; Dr. H. Yang, Oregon Graduate Institute of Science and Technology; Dr. Scott Douglas, University of Utah; Dr. Pierre Comon,

Thomson-Sintra Asm., France; Dr. Terrence Sejnowski, Salk Institute; Dr. Harris Drucker, Monmouth College; Dr. Nathan Intrator, Tel Aviv University, Israel; Dr. Vladimir Vapnik, AT&T Research, NJ; Dr. Teuvo Kohonen, Helsinki University of Technology, Finland; Dr. Vladimir Cherkassky, University of Minnesota; Dr. Sebastian Seung, AT&T Research, NJ; Dr. Steve Luttrell, DERA, Great Malvern, United Kingdom; Dr. David Lowe, Aston University, United Kingdom; Dr. N. Ansari, New Jersey Institute of Technology; Dr. Danil Prokhorov, Ford Research Laboratory, Dearborn, Mich.; Dr. Shigeru Katagiri, ATR Human Information Processing Research Lab, Japan; Dr. James Anderson, Brown University; Dr. Irwin Sandberg, University of Texas at Austin; Dr. Thomas Cover, Stanford University; Dr. Walter Freeman, University of California at Berkeley; Dr. Charles Micchelli, IBM Research, Yorktown Heights; Dr. Kari Torkkola, Motorola Phoenix Corp.; Dr. Andreas Andreou, Johns Hopkins University; Dr. Martin Beckerman, Oak Ridge National Laboratory; and Dr. Thomas Anastasio, University of Illinois, Urbana.

I am deeply indebted to my graduate student Hugh Pasika for performing many of the MATLAB experiments in the book, and for preparing the Web site for the book. The help received from my graduate student Himesh Madhuranath, Dr. Sadasivan Puthusserypady, Dr. J. Nie, Dr. Paul Yee, and Mr. Gint Puskorius (Ford Research) in performing five of the experiments is much appreciated.

I am most grateful to Hugh Pasika for proofreading the entire book. In this regard, I also thank Dr. Robert Dony (University of Guelph), Dr. Stefan Kremer (University of Guelph), and Dr. Sadasivan Puthusserypaddy for proofreading selected chapters of the book.

I am grateful to my publisher Tom Robbins and editor Alice Dworkin for their full support and encouragement. The careful copy editing of the manuscript by Julie Hollist is much appreciated. I would like to thank the tireless efforts of Jennifer Maughan and the staff of WestWords Inc. in Logan, Utah in the production of the book.

I wish to record my deep gratitude to Brigitte Maier, Thode Library, McMaster University, for her untiring effort to search for and find very difficult references that have made the bibliography all the more complete. The help of Science and Engineering Librarian Peggy Findlay and Reference Librarian Regina Bendig is also much appreciated.

Last but by no means least, I am most grateful to my secretary Lola Brooks for typing so many different versions of the manuscript. Without her dedicated help, the writing of this book and its production would have taken a great deal longer.



Abbreviations and Symbols

ABBREVIATIONS

AI	artificial intelligence
APEX	adaptive principal components extraction
AR	autoregressive
BBTT	back propagation through time
BM	Boltzmann machine
BP	back propagation
b/s	bits per second
BOSS	bounded, one-sided saturation
BSB	brain-state-in-a-box
BSS	Blind source (signal) separation
CART	classification and regression tree
cmm	correlation matrix memory
CV	cross-validation
DEKF	decoupled extended Kalman filter
DFA	deterministic finite-state automata
DSP	digital signal processor
EKF	extended Kalman filter
EM	expectation-maximization
FIR	finite-duration impulse response
FM	frequency-modulated (signal)
GEKF	global extended Kalman filter
GCV	generalized cross-validation
GHA	generalized Hebbian algorithm
GSLC	generalized sidelobe canceler

HME	hierarchical mixture of experts
HMM	hidden Markov model
Hz	hertz
ICA	independent components analysis
Infomax	maximum mutual information
KR	kernel regression
LMS	least-mean-square
LR	likelihood ratio
LTP	long-term potentiation
LTD	long-term depression
LR	likelihood ratio
LVQ	learning vector quantization
MCA	minor components analysis
MDL	minimum description length
ME	mixture of experts
MFT	mean-field theory
MIMO	multiple input–multiple output
ML	maximum likelihood
MLP	multilayer perceptron
MRAC	model reference adaptive control
NARMA	nonlinear autoregressive moving average
NARX	nonlinear autoregressive with exogenous inputs
NDP	neuron-dynamic programming
NW	Nadaraya–Watson (estimator)
NWKR	Nadaraya–Watson kernel regression
OBD	optimal brain damage
OBS	optimal brain surgeon
OCR	optical character recognition
ODE	ordinary differential equation
PAC	probably approximately correct
PCA	principal components analysis
pdf	probability density function
pmf	probability mass function
RBF	radial basis function
RMLP	recurrent multilayer perceptron
RTRL	real-time recurrent learning
SIMO	single input–multiple output
SISO	single input–single output
SNR	signal-to-noise ratio
SOM	self-organizing map

SRN	simple recurrent network (also referred to as Elman's recurrent network)
SVD	singular value decomposition
SVM	support vector machine
TDNN	time-delay neural network
TLFN	time lagged feedforward network
VC	Vapnik–Chervononkis (dimension)
VLSI	very-large-scale integration
XOR	exclusive OR

IMPORTANT SYMBOLS

a	action
$\mathbf{a}^T \mathbf{b}$	inner product of vectors \mathbf{a} and \mathbf{b}
$\mathbf{a} \mathbf{b}^T$	outer product of vectors \mathbf{a} and \mathbf{b}
$\binom{l}{m}$	binomial coefficient
$A \cup B$	unions of A and B
B	inverse of temperature
b_k	bias applied to neuron k
$\cos(\mathbf{a}, \mathbf{b})$	cosine of the angle between vectors \mathbf{a} and \mathbf{b}
D	depth of memory
$D_{f g}$	Kullback–Leibler divergence between probability density functions f and g
$\tilde{\mathbf{D}}$	adjoint of operator \mathbf{D}
E	energy function
E_i	energy of state i in statistical mechanics
\bar{E}	statistical expectation operator
$\langle E \rangle$	average energy
erf	error function
erfc	complimentary error function
exp	exponential
\mathcal{E}_{av}	average squared error or sum of squared errors
$\mathcal{E}(n)$	instantaneous value of the sum of squared errors
\mathcal{E}_{total}	total sum of error squares
F	free energy
$f_{\mathbf{x}}(\mathbf{x})$	probability density function of random vector \mathbf{X}
\mathcal{F}^*	subset (network) with the smallest minimum empirical risk
\mathbf{H}	Hessian matrix
\mathbf{H}^{-1}	inverse of matrix \mathbf{H}
i	square root of -1 , also denoted by j
\mathbf{I}	identity matrix
\mathbf{I}	Fisher's information matrix
J	mean-square error