Artificial Neural Networks: Concepts and Theory

Pankaj Mehra and Benjamin W. Wah

IEEE Computer Society Press
Los Alamitos, California
Washington • Brussels • Tokyo

IEEE COMPUTER SOCIETY PRESS TUTORIAL

Library of Congress Cataloging-in-Publication Data

Artificial neural networks: concepts and theory / (compiled by)
Pankaj Mehra and Benjamin W. Wah.
p. cm. -- (IEEE Computer Society Press Tutorial)
Includes bibliographical references.
ISBN 0-8186-8997-8 (case) -- ISBN 0-8186-5997-1 (m/f).
1. Neural networks (Computer science)
I. Mehra, Pankaj, 1964 - II. Wah, Benjamin W. III. Series
QA76.87.A74 1992
006.3--dc20 91-46288
CIP

Published by the
IEEE Computer Society Press
10662 Los Vaqueros Circle
PO Box 3014
Los Alamitos, CA 90720-1264

IEEE Computer Society Press Order Number 1997
Library of Congress Number 91-46288
IEEE Catalog Number 92EH0354-1
ISBN 0-8186-5997-1 (microfiche)
ISBN 0-8186-8997-8 (case)

Additional copies can be ordered from

IEEE Computer Society Press
Customer Service Center
10662 Los Vaqueros Circle
PO Box 3014
Los Alamitos, CA 90720-1264

IEEE Service Center
445 Hoes Lane
PO Box 1331
Piscataway, NJ 08855-1331

IEEE Computer Society
13, avenue de l'Aquilon
B-1200 Brussels
BELGIUM

IEEE Computer Society
Ooshima Building
2-19-1 Minami-Aoyama
Minato-ku, Tokyo 107
JAPAN

Editorial production: Penny Storms
Copy Editor: Phyllis Walker
Cover Design: Joe Daigle
Printed in the United States of America by Victor Graphics, Inc.

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

Acknowledgments

We are indebted to the reviewers who suggested numerous improvements to our original manuscript. We would like to thank Drs. Rao Vemuri and Joydeep Ghosh, for their editorial assistance and for working closely with us in improving the quality of our presentation, and Drs. Jose Fortes and Manoel Tenorio, for working jointly with us on earlier drafts of this tutorial. We especially thank the following persons for their comments and criticisms on earlier drafts: Andrew Barron, Subutai Ahmad, Munindar Singh, Mark Gooley, Kevin Buescher, and Chris Matheus. Mr. Mehra wishes to thank his wife Ranjana, for egging him on; his parents Kamal and Motia, for instilling in him the quest to learn; and Professors Larry Rendell, Andrew Barron, P. R. Kumar, and Kanad Biswas, for teaching him the basics of machine learning and neural networks. Dr. Wah wishes to thank his wife Christine, for bearing with him in the preparation for presenting this tutorial in numerous conferences, and Drs. Henry Lum and Bernard Chern, for various discussions in developing this tutorial. Finally, we are both grateful for the support of the National Aeronautics and Space Administration, under Contract NCC 2-481, and the National Science Foundation, under grant MIP 88-10584, during the preparation of this tutorial.

Pankaj Mehra and Benjamin W. Wah
June 21, 1992

Preface

Artificial neural networks (ANNs) are computational structures modeled on biological processes. In recent years, ANNs have been used for implementing nonlinear controllers, content-addressable memory, optimization, constraint satisfaction, pattern classification, and dimensionality reduction. They are being promoted for their robustness, massive parallelism, and ability to learn.

One of the major problems in studying ANNs is that the literature in this area is scattered over a vast number of publications spanning several somewhat unrelated disciplines. Computer scientists and information theorists, treating ANNs as learning systems, design and analyze ANN algorithms for generality, efficiency, accuracy, and robustness. Psychologists and cognitive scientists study ANNs as abstract models of human and animal nervous systems. Electrical engineers and physicists address the feasibility of implementing ANNs in VLSI and optics, while computer engineers study the simulation of ANNs on parallel computers. Application engineers focus on using ANNs in innovative applications, such as robotics, speech recognition, and image understanding. Finally, mathematicians view ANNs as abstract objects (functions and dynamic systems) and study their static and dynamic properties.

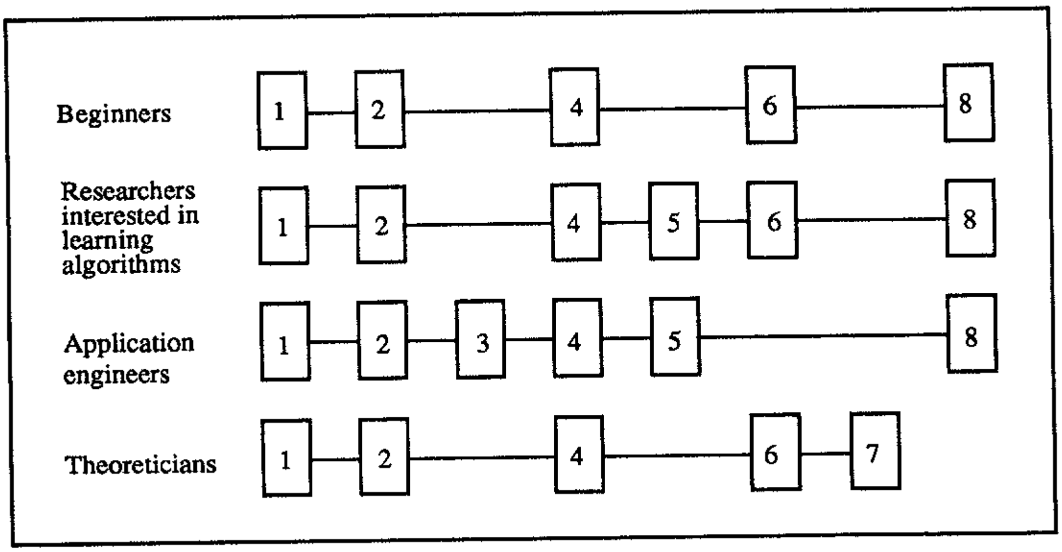
The literature on representation and learning using ANNs spans many diverse and somewhat unrelated disciplines. The vast body of ANN research, as well as the diversity of its sources, makes it difficult for novices to learn about ANNs and for researchers to keep pace with current developments. Someone attempting to start learning about ANNs may be overwhelmed by the sheer volume of recent papers; unfamiliarity with ANN terminology and certain commonly used concepts can only add to this problem. Consequently, it is difficult to get a holistic picture of ANNs and to combine and apply results from diverse sources to practical problems.

Our goal was to make the current literature on ANNs accessible to students, academicians, engineers, and other professionals who want to learn about the field, as well as to researchers, who can use this tutorial to become informed about current research. For the beginner, we introduce terminology and classification, guide the reader through key recent papers, and provide pointers for further reading. For the active researcher, we present a reference collection of key papers and an up-to-date bibliography of current literature. For the engineer and the practitioner, we introduce the theory behind the tools for designing and analyzing ANNs, and we cite the more advanced papers for in-depth reading on specific aspects. We focus on basic concepts, algorithms, and theoretical results, and we emphasize the interplay between abstract theoretical issues and practical design issues, thus reflecting the current trends of ANN research. We neither discuss biological aspects of neural networks nor evaluate ANNs as models of cognition. Even though we do not cover examples of specific applications of ANNs in such fields as robotics, image understanding, and optimization, we do discuss the theory underlying the use of ANNs in these applications. Finally, we do not discuss the implementation of ANNs in hardware (semiconductor or optical), software simulation, or commercial implementations.

The tutorial is divided into eight chapters. Chapter 1 introduces the basic terminology of ANNs, identifies their characteristic traits, and presents various classifications of ANN research. Chapter 2 describes the variety of ANN structures available today and examines further the basic components of an ANN model. Chapter 3 shows how knowledge may be represented in ANNs, illustrating how ANNs may carry out intelligent reasoning and problem-solving tasks. The next two chapters cover learning algorithms for ANNs. Chapter 4 further develops the classification scheme and terminology introduced in Chapter 1 and covers the most popular form of ANN learning: one that uses deterministic rules to perform supervised learning. Other learning rules are covered in Chapter 5. The next two chapters cover the theory of ANNs. First, in Chapter 6, we study ANNs without feedback, which can be analyzed using tools from approximation theory and information theory. Then, in Chapter 7, we study ANNs with feedback, whose analysis requires tools from the theory of differential equations and dynamic systems. Finally, Chapter 8 is devoted to empirical studies of ANN behavior, including observation and characterization of the representation, problem-solving, and learning abilities of ANNs. Each chapter first introduces the terminology for — and provides an overview of — the topic addressed, and then includes a small number of key papers from recent literature on that topic. Size limitations prohibit the inclusion of all important papers in the field; we regret any omissions we have made. An extensive list of references at the end of each chapter points to papers relevant to that chapter's topic. Reference citations for papers that appear as reprints in that chapter are marked with an asterisk.

In choosing papers to include in this tutorial, we attempted to complement existing tutorials and edited books, overlapping with these as little as possible. Chapters are divided into sections; the grouping of papers included in each chapter parallels this organization. An attempt was made to include one paper per section. In our discussion of included papers, we provide cross-references to relevant sections of both the chapter introductions and of other included papers.

The flowcharts presented in the figure below suggest orders in which readers approaching ANNs from various perspectives might choose to read the chapters of this tutorial. Chapters 1 and 2 contain introductory discussions of ANNs and illustrate various types of ANN structures; while those familiar with ANNs may just skim through these discussions, others will need to read them before progressing to other parts of the tutorial. Beginners may follow up Chapters 1 and 2 by reading the introductory materials on learning in Chapter 4, on theory in Chapter 6, and on experimentation in Chapter 8. Researchers interested in learning algorithms may wish to focus first on the discussions of algorithms and theory of learning in Chapters 4, 5, and 6 and then on the discussion of experimental issues in Chapter 8. Application engineers may skip over Chapters 6 and 7, which contain discussion of theoretical aspects; in Chapter 3, they will find an important discussion of knowledge representation using ANNs. Finally, theoreticians may wish to follow up the introductory material of Chapters 1 and 2 with an introduction to learning in Chapter 4, then focusing on Chapters 6 and 7, which address theoretical issues.



Flowcharts — from various reader perspectives — showing the suggested order for reading chapters of this tutorial.

Table of Contents