Neural Computing

Theory and Practice

Philip D. Wasserman

Neural Computing

Theory and Practice

Philip D. Wasserman

ANZA Research, Inc.

VNR VAN NOSTRAND REINHOLD
New York

Library of Congress Catalog Card Number 88-34842
ISBN 0-442-20743-3

Printed in the United States of America

Van Nostrand Reinhold
115 Fifth Avenue
New York, New York 10003

Van Nostrand Reinhold International Company Limited
11 New Fetter Lane
London EC4P 4EE, England

Van Nostrand Reinhold
480 La Trobe Street
Melbourne, Victoria 3000, Australia

Nelson Canada
1120 Birchmount Road
Scarborough, Ontario M1K 5G4, Canada

16 15 14 13 12 11 10 9 8 7 6 5 4 3 2

Library of Congress Cataloging-in-Publication Data
Wasserman, Philip D., 1937–
Neural computing : theory and practice / Philip D. Wasserman.
p. cm.
Includes bibliographies and index.
ISBN 0-442-20743-3
1. Neural computers. I. Title.
QA76.5.W353 1989 88-34842
006.3—dc19 CIP

Preface

What are artificial neural networks? What can they do? How do they work? How can I use them? These and many similar questions are being asked by professionals from a wide variety of disciplines. Finding comprehensible answers has been difficult. University courses are few, seminars are expensive, and the literature is extensive and specialized. The several excellent books in print can prove daunting. Often expressed in technical jargon, many of the treatments assume a facility with branches of advanced mathematics that are seldom used in other specialties.

This book provides a systematic entry path for the professional who has not specialized in mathematical analysis. All of the important concepts are first expressed in ordinary English. Informal mathematical treatments are included when they clarify the explanation. Complicated derivations and proofs are placed at the end of chapters and references to other works are regularly provided. These references constitute an extensive bibliography of important writings in specific areas applicable to artificial neural networks. This multilevel approach not only provides the reader with an overview of artificial neural networks but also permits the serious student an in-depth exploration of the subject.

Every effort has been made to produce a book that is easily understood without oversimplification of the material. Readers who go on to more theoretical studies should not need to unlearn anything presented here. When simplifications are employed, they are labeled as such and references point to more detailed treatments.

This book need not be read from cover to cover. Each chapter is intended to be self-contained, assuming familiarity only with the topics of Chapters 1 and 2. While this implies a certain amount of repetition, most readers should not find it onerous.

Practicality has been a primary objective. If the chapters are studied carefully, it should be possible to implement most of the networks on a general-purpose computer. The reader is urged to do so; no other method will produce the same depth of understanding.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my wife, Sarah, for her encouragement and tolerance during the months I spent in the company of my word processor.

Also, I would like to thank my friends and colleagues, who gave so generously of their time and knowledge, corrected my errors, and created an environment in which ideas developed rapidly. I would like to extend my special appreciation to Dr. Surapol Dasananda, Santa Clara University; Dr. Elizabeth Center, College of Notre Dame; Dr. Peter Rowe, College of Notre Dame; Charles Rockwell, Microlog Corp.; Tom Schwartz, The Schwartz Associates; Dennis Reinhardt, Dair Corp.; Coe Miles-Schlichting; and Douglas Marquardt. Thanks also are due to Kyla Carlson and Nang Cao for their help in preparing the illustrations.

I must, of course, take the blame for any residual errors; they couldn't watch me every minute.

Contents

Introduction

WHY ARTIFICIAL NEURAL NETWORKS?

After two decades of near eclipse, interest in artificial neural networks has grown rapidly over the past few years. Professionals from such diverse fields as engineering, philosophy, physiology, and psychology are intrigued by the potential offered by this technology and are seeking applications within their disciplines.

This resurgence of interest has been fired by both theoretical and application successes. Suddenly, it appears possible to apply computation to realms previously restricted to human intelligence; to make machines that learn and remember in ways that bear a striking resemblance to human mental processes; and to give a new and significant meaning to the much-abused term *artificial intelligence.*

CHARACTERISTICS OF ARTIFICIAL NEURAL NETWORKS

Artificial neural networks are biologically inspired; that is, they are composed of elements that perform in a manner that is analogous to the most elementary functions of the biological neuron. These elements are then organized in a way that may (or may not) be related to the anatomy of the brain. Despite this superficial resemblance, artificial neural networks exhibit a surprising number of the brain's characteristics. For example, they learn from experience, generalize from previous examples to new ones, and abstract essential characteristics from inputs containing irrelevant data.

Despite these functional similarities, not even the most optimistic advocate will suggest that artificial neural networks will soon duplicate the functions of the human brain. The actual "intelligence" exhibited by the most sophisticated artificial neural networks is below the level of a tapeworm; enthusiasm must be tempered by current reality. It is, however, equally incorrect to ignore the surprisingly brainlike performance of certain artificial neural networks. These abilities, however limited they are today, hint that a deep understanding of human intelligence may lie close at hand, and along with it a host of revolutionary applications.

Learning

Artificial neural networks can modify their behavior in response to their environment. This factor, more than any other, is responsible for the interest they have received. Shown a set of inputs (perhaps with desired outputs), they self-adjust to produce consistent responses. A wide variety of training algorithms has been developed, each with its own strengths and weaknesses. As we point out later in this volume, there are important questions yet to be answered regarding what things a network can be trained to do, and how the training should be performed.

Generalization

Once trained, a network's response can be, to a degree, insensitive to minor variations in its input. This ability to see through noise and distortion to the pattern that lies within is vital to pattern recognition in a real-world environment. Overcoming the literal-mindedness of the conventional computer, it produces a system that can deal with the imperfect world in which we live. It is important to note that the artificial neural network generalizes automatically as a result of its structure, not by using human intelligence embedded in the form of ad hoc computer programs.

Abstraction

Some artificial neural networks are capable of abstracting the essence of a set of inputs. For example, a network can be trained on a sequence of distorted versions of the letter A. After adequate training, application of such a distorted example will cause the network to produce a perfectly formed letter. In one sense, it has learned to produce something that it has never seen before.

This ability to extract an ideal from imperfect inputs raises interesting philosophical issues; it is reminiscent of the concept of ideals found in Plato's *Republic*. In any case, extracting idealized prototypes is a highly useful ability in humans; it seems that now we may share it with artificial neural networks.

Applicability

Artificial neural networks are not a panacea. They are clearly unsuited to such tasks as calculating the payroll. It appears that they will, however, become the preferred technique for a large class of pattern-recognition tasks that conventional computers do poorly, if at all.

HISTORICAL PERSPECTIVE

Humans have always wondered about their own thoughts. This self-referential inquiry, the mind thinking of itself, may be a uniquely human characteristic. Speculations on the nature of thought abound, ranging from the spiritual to the anatomical. With philosophers and theologians often opposing the opinions of physiologists and anatomists, the questions have been hotly debated to little avail, as the subject is notoriously difficult to study. Those relying on introspection and speculation have arrived at conclusions that lack the rigor demanded by the physical sciences. Experimenters have found the brain and nervous system to be difficult to observe and perplexing in organization. In short, the powerful methods of scientific inquiry that have changed our view of physical reality have been slow in finding application to the understanding of humans themselves.

Neurobiologists and neuroanatomists have made substantial progress. Painstakingly mapping out the structure and function of the human nervous system, they came to understand much of the brain's "wiring," but little of its operation. As their knowledge grew, the complexity was found to be staggering. Hundreds of billions of neurons, each connecting to hundreds or thousands of others, comprise a system that dwarfs our most ambitious dreams of supercomputers. Nevertheless, the brain is gradually yielding its secrets to one of humankind's most sustained and ambitious inquiries.

The improved understanding of the functioning of the neuron and the pattern of its interconnections has allowed researchers to produce mathematical models to test their theories. Experiments can now be conducted on digital computers without involving human or animal subjects, thereby solving many practical and ethical problems. From early work it became apparent that these models not only mimicked functions of the brain, but that they were capable of performing useful functions in their own right. Hence, two mutually reinforcing objectives of neural modeling were defined and remain today: first, to understand the physiological and psychological functioning of the human neural system; and second, to produce computational systems (artificial neural networks) that perform brainlike functions. It is the latter objective that is the major focus of this book.

Along with the progress in neuroanatomy and neurophysiology, psychologists were developing models of human learning. One such model, which has proved most fruitful, was that of D. O. Hebb, who in 1949 proposed a learning law that became the starting point for artificial neural network training algorithms. Augmented today by many other methods, it showed scientists of that era how a network of neurons could exhibit learning behavior.

In the 1950s and 1960s, a group of researchers combined these biological and psychological insights to produce the first artificial neural networks. Initially implemented as electronic circuits, they were later converted to the more flexible medium of computer simulation, the most common realization today. Early successes produced a burst of activity and optimism. Marvin Minsky, Frank Rosenblatt, Bernard Widrow, and others developed networks consisting of a single layer of artificial neurons. Often called perceptrons, they were applied to such diverse problems as weather prediction, electrocardiogram analysis, and artificial vision. It seemed for a time that the key to intelligence had been found; reproducing the human brain was only a matter of constructing a large enough network.

This illusion was soon dispelled. Networks failed to solve problems superficially similar to those they had been successful in solving. These unexplained failures launched a period of intense analysis. Marvin Minsky, carefully applying mathematical techniques, developed rigorous theorems regarding network operation. His research led to the publication of the book *Perceptrons* (Minsky and Papert 1969), in which he and Seymore Papert proved that the single-layer networks then in use were theoretically incapable of solving many simple problems, including the function performed by a simple exclusive-or gate. Nor was Minsky optimistic about the potential for progress:

> The Perceptron has shown itself worthy of study despite (and even because of!) its severe limitations. It has many features that attract attention: its linearity; its intriguing learning theorem; its clear paradigmatic simplicity as a kind of parallel computation. There is no reason to suppose that any of these virtues carry over to the many-layered version. Nevertheless, we consider it to be an important research problem to elucidate (or reject) our intuitive judgment that the extension is sterile.
>
> Perhaps some powerful convergence theorem will be discovered, or some profound reason for the failure to produce an interesting "learning theorem" for the multilayered machine will be found. (pp. 231–32)

Minsky's brilliance, rigor, and prestige gave the book great credibility: its conclusions were unassailable. Discouraged researchers left the field for areas of greater promise, government agencies redirected their funding, and artificial neural networks lapsed into obscurity for nearly two decades.

Nevertheless, a few dedicated scientists such as Teuvo Kohonen, Stephen Grossberg, and James Anderson continued their efforts. Often underfunded and unappreciated, some researchers had difficulty finding publishers; hence, research published during the

1970s and early 1980s is found scattered among a wide variety of journals, some of which are rather obscure. Gradually, a theoretical foundation emerged, upon which the more powerful multilayer networks of today are being constructed. Minsky's appraisal has proven excessively pessimistic; networks are now routinely solving many of the problems that he posed in his book.

In the past few years, theory has been translated into application, and new corporations dedicated to the commercialization of the technology have appeared. There has been an explosive increase in the amount of research activity. With four major conventions in 1987 in the field of artificial neural networks, and over 500 technical papers published, the growth rate has been phenomenal.

The lesson to be learned from this history is found in Clark's law. Propounded by the writer and scientist Arthur C. Clark, it states in effect that if a respected senior scientist says a thing can be done, he or she is almost always correct; if the scientist says it cannot be done, he or she is almost always wrong. The history of science is a chronicle of mistakes and partial truths. Today's dogma becomes tomorrow's rubbish. Unquestioning acceptance of "facts," whatever the source, can cripple scientific inquiry. From one point of view, Minsky's excellent scientific work led to an unfortunate hiatus in the progress of artificial neural networks. There is no doubt, however, that the field had been plagued by unsupported optimism and an inadequate theoretical basis. It may be that the shock provided by *Perceptrons* allowed a period for the necessary maturation of the field.

ARTIFICIAL NEURAL NETWORKS TODAY

There have been many impressive demonstrations of artificial neural network capabilities: a network has been trained to convert text to phonetic representations, which were then converted to speech by other means (Sejnowsky and Rosenberg 1987); another network can recognize handwritten characters (Burr 1987); and a neural network–based image-compression system has been devised (Cottrell, Munro, and Zipser 1987). These all use the backpropagation network, perhaps the most successful of the current algorithms. Backpropagation, invented independently in three separate research efforts (Werbos 1974; Parker 1982; and Rumelhart, Hinton, and Williams 1986), provides a systematic means for training multilayer networks, thereby overcoming limitations presented by Minsky.

As we point out in the chapters that follow, backpropagation is not without its problems. First, there is no guarantee that the network can be trained in a finite amount of time. Many training efforts fail after consuming large amounts of computer time. When this happens, the training attempt must be repeated—with no certainty that the results will be any better. Also, there is no assurance that the network will train to the best configuration possible. So-called local minima can trap the training algorithm in an inferior solution.

Many other network algorithms have been developed that have specific advantages; several of these are discussed in the chapters that follow. It should be emphasized that none of today's networks represents a panacea; all of them suffer from limitations in their ability to learn and recall.

We are presented with a field having demonstrated performance, unique potential, many limitations, and a host of unanswered questions. It is a situation calling for optimism tempered with caution. Authors tend to publish their successes and give their failures little publicity, thereby creating an impression that may not be realistic. Those seeking venture capital to start new firms must present a convincing projection of substantial accomplishments and profits. There exists, therefore, a substantial danger that artificial neural networks will be oversold before their time, promising performance without the capability for delivery. If this happens, the entire field could suffer a loss of credibility, possibly relapsing into the Dark Ages of the 1970s. Much solid work is required to improve existing networks. New techniques must be developed, existing methods strengthened, and the theoretical foundation broadened before this field can realize its full potential.

PROSPECTS FOR THE FUTURE

Artificial neural networks have been proposed for tasks ranging from battlefield management to minding the baby. Potential applications are those where human intelligence functions effortlessly and conventional computation has proven cumbersome or inadequate. This application class is at least as large as that serviced by conventional computation, and the vision arises of artificial neural networks taking their place alongside of conventional computation as an adjunct of equal size and importance. This will happen only if fundamental research yields results at a rapid rate, as today's theoretical foundations are inadequate to support such projections.

Artificial Neural Networks and Expert Systems

The field of artificial intelligence has been dominated in recent years by the logical- and symbol-manipulation disciplines. For example, expert systems have been widely acclaimed and have achieved many notable successes—as well as many failures. Some say that artificial neural networks will replace current artificial intelligence, but there are many indications that the two will coexist and be combined into systems in which each technique performs the tasks for which it is best suited.

This viewpoint is supported by the way that humans operate in the world. Activities requiring rapid responses are governed by pattern recognition. Since actions are produced rapidly and without conscious effort, this mode of operation is essential for the quick, spontaneous responses needed to survive in a hostile environment. Consider the consequences if our ancestors had to reason out the correct response to a leaping carnivore!

When our pattern-recognition system fails to produce an unambiguous interpretation (and when time permits), the matter is referred to the higher mental functions. These may require more information and certainly more time, but the quality of the resulting decisions can be superior.

One can envision an artificial system that mimics this division of labor. An artificial neural network would produce an appropriate response to its environment under most circumstances. Because such networks can indicate the confidence level associated with each decision, it would "know that it did not know," and would refer that case to an expert system for resolution. The decisions made at this higher level would be concrete and logical, but might require the gathering of additional facts before a conclusion could be reached. The combination of the two systems would be more robust than either acting alone, and it would follow the highly successful model provided by biological evolution.

Reliability Considerations

Before artificial neural networks can be applied where human life or valuable assets are at stake, questions regarding their reliability must be resolved.

Like the humans whose brain structure they mimic, artificial neural networks retain a degree of unpredictability. Unless every possible input is tried, there is no way to be certain of the precise output. In a large network such exhaustive testing is impractical and statistical estimates of performance must suffice. Under some circumstances this is intolerable. For example, what is an acceptable error rate for a network controlling a space defense system? Most people would say that any error is intolerable; it might result in unthinkable death and destruction. This attitude is not changed by the fact that a human in the same situation might also make mistakes.

The problem lies in the expectation that computers are absolutely error free. Because artificial neural networks will sometimes make errors even when they are functioning correctly, many feel that this translates into unreliability, a characteristic we have found unacceptable in our machines.

A related difficulty lies in the inability of traditional artificial neural networks to "explain" how they solve problems. The internal representations that result from training are often so complex as to defy analysis in all but the most trivial cases. This is closely related to our inability to explain how we recognize a person despite differences in distance, angle, illumination, and the passage of years. An expert system can trace back through its own reasoning process so that a human can check it for reasonableness. Incorporation of this ability into artificial neural networks has been reported (Gallant 1988) and its development may have an important effect upon the acceptability of these systems.

SUMMARY

Artificial neural networks represent a major extension of computation. They promise human-made devices that perform functions heretofore reserved for human beings. Dull, repetitive, or dangerous tasks can be performed by machines and entirely new applications will arise as the technology matures.

The theoretical foundations of artificial neural networks are expanding rapidly, but they are currently inadequate to support the more optimistic projections. Viewed historically, theory has developed faster than pessimists had projected and slower than optimists had hoped, a typical situation. Today's surge of interest has set thousands of researchers to work in the field. It is reasonable to expect a rapid increase in our understanding of artificial neural networks leading to improved network paradigms and a host of application opportunities.

References

Burr, D. J. 1987. Experiments with a connectionist text reader. In *Proceedings of the First International Conference on Neural Networks,* eds. M. Caudill and C. Butler, vol. 4, pp. 717–24. San Diego, CA: SOS Printing.

Cottrell, G. W., Munro, P., and Zipser, D. 1987. Image compression by backpropagation: An example of extensional programming. *Advances in cognitive science* (vol. 3). Norwood, NJ: Ablex.

Gallant, S. I. 1988. Connectionist expert systems. *Communications of the ACM* 31:152–69.

Minsky, M., and Papert, S. 1969. *Perceptrons.* Cambridge, MA: MIT Press.

Parker, D. B. 1982. Learning-logic. Invention Report, S81-64, File 1. Office of Technology Licensing, Stanford University.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. 1986. Learning internal representations by error propagation. In *Parallel distributed processing*, vol. 1, pp. 318–62. Cambridge, MA: MIT Press.

Sejnowski, T. J., and Rosenberg, C. R. 1987. Parallel networks that learn to pronounce English text. *Complex Systems* 3:145–68.

Werbos, P. J. 1974. *Beyond regression: New tools for prediction and analysis in the behavioral sciences.* Masters thesis, Harvard University.