Lecture Notes in Computer Science 931
Edited by G. Goos, J. Hartmanis and J. van Leeuwen

Advisory Board: W. Brauer D. Gries J. Stoer

P.J. Braspenning F. Thuijsman
A.J.M.M. Weijters (Eds.)

Artificial Neural Networks

An Introduction to ANN Theory and Practice

Springer

Volume Editors

P.J. Braspenning, Department of Computer Science
F. Thuijsman, Department of Mathematics
A.J.M.M. Weijters, Department of Computer Science
University of Limburg, P.O. Box 616
6200 MD Maastricht, The Netherlands

CR Subject Classification (1991): F1.1, I.2.6, G.1.6, I.5.1, J.1, J.2, J.6

1991 Mathematics Subject Classification: 92B20, 94C15, 68T05, 90C90

ISBN 3-540-59488-4 Springer-Verlag Berlin Heidelberg New York

CIP data applied for

Typesetting: Camera-ready by author
SPIN: 10486258 06/3142-543210 - Printed on acid-free paper

Preface

This book is the result of a concerted action by the departments of Computer Science and Mathematics of the University of Limburg (Maastricht, The Netherlands) to develop a collection of lectures, specifically dedicated to *informing the industrial world* about the *potential* of using neural networks. For this reason, both departments had worked together within an *NN working group* to set up an Autumn School for Neural Networks, which was held in 1990 in Maastricht. Participants came from different quarters within government, industry and small and medium-sized companies, and insurance and banking institutes. However, the participants were not *arbitrarily chosen* workers within those quarters.

The target group of people addressed by the Neural Network School were *technical managers, consultants, research associates, and software developers* at the high end of the spectrum whose employers expected innovative applications of new technologies within their own (industrial) setting. Hence, in our view the target group consisted of people with a *reasonable level* of formal education and, specifically, some basic background in mathematics and computer science. Having this group in mind, the contributions of this book were set. Hence the prerequisites for a fruitful understanding of the material are set. Nevertheless, some more specific knowledge of mathematics and/or computer science may be required at a few places.

The aim of this book is not to offer a systematic exposition of all kinds of neural networks or a bunch of most often used networks. Rather, the idea was to focus on two generic application domains, namely *control* and *optimization*, and use these application domains to illustrate the concrete use of different kinds of neural networks. Put otherwise, these application domains were used to cluster and direct the NN School lectures to be particularly illustrative for how to apply different kinds of neural network architecture. In this way, we hoped to serve the needs of the participants, both regarding their need to (theoretically) *understand the functioning* of any particular network and regarding their need to really see a *demonstrative example* application.

After the NN School was held we used the feedback from the participants to *update* and *elaborate* the course materials into a set of papers which together comprise this book. That is, the book is a compilation, not of the original course materials, but of carefully re-worked original papers. However, it should not be seen as detailing the newest developments within the rapidly evolving *scientific* discipline of neural networks. As explained already, this was not the goal of our efforts.

What can the reader expect of this book? First, it gives a *representative* bunch of neural network architectures which have found widespread application. The level of exposition is such that the functioning of these neural networks can be understood, and many times their functioning is also dealt with in a more analytical fashion. Secondly, quite a few applications are described for which a *particular* neural network architecture has been chosen. This choice is not always

based on purely objective criteria, because the field of neural networks is still of a rather experimental nature. However, where possible the actual choice of architecture is reasoned. In fact, one contributing paper in this book is exclusively dedicated to *making a choice* about the neural network architecture to use *for a particular task* within an actual application domain. Thirdly, reading the book as a whole certainly *stimulates one's curiosity* (and therefore one's innovativeness) *about the applicability* of neural networks within one's own field of work. Therefore, the team of authors considers itself to have been successful if many readers, after reading this book, seriously consider applying some neural network technology to the problem at hand, whether it is for a *classification/recognition* task, a *control* task, or a complex multiple *constraint satisfaction* task.

Of course, the ordering of the papers in this book is not arbitrary. It constitutes *the route* which we think to be most profitable for the serious reader. However, depending on the reader's pre-existing background knowledge about neural networks, we do not object at all to a reader who wants to dive into particular papers, especially because all papers are *sufficiently self-contained.* Nevertheless, we would like to finish this introduction with an outline of the route of papers in this book.

The first paper, by Braspenning, gives a general chararcterization of neural networks and puts the contributions to the book in perspective. In the contribution by Weijters and Hoppenbrouwers the *back-propagation network* is discussed. This is probably the most widespread and most popular architecture. The paper by Henseler addresses this architecture again, but in a more formal way and applied to a *robot control* task. The paper by Peters treats the forerunner of back-propagation networks, namely *perceptrons*, and analyzes mathematically their advantages/disadvantages. The contribution by Vrieze gives a basic treatment of another architecture, the *Kohonen network*, and analyzes basic expectations of what it does and can do for a number of application tasks. The paper by Postma and Hudson again introduces an architecture, namely *adaptive resonance networks*, and discusses a number of variants. The paper by Spieksma treats a neural network architecture which is inspired by physical phenomena treated by statistical mechanics, namely the *Boltzmann machine* architecture, and applies it to a combinatorial optimization problem. The contribution by Lenting discusses the same architecture, but now from the perspective of how to map (or *represent*) *a particular problem* on (with) such an architecture; a topic which, in fact, deserves careful attention with any of the architectures. The paper by Postma introduces another architecture inspired by a physical theory, namely the *Hopfield-Tank* network, but its main gist is to show how neural networks can help in solving optimization problems. The contribution by Crama, Kolen, and Pesch addresses a wide range of *combinatorial optimization* approaches including a neural network approach like the Boltzmann machine and a genetic algorithm inspired by a Darwinian framework. The paper by Boekhoudt turns to *process identification and control*, which is another important generic application domain of neural networks, and some already introduced NN architectures are evaluated regarding their promising use. The paper by Van Luenen again

deals with control tasks and discusses the neural network design and *learning* strategies appropriate for these tasks with a quite illustrative example application: the inverted pendulum. The paper by Cardon and Hoogstraten is written from the perspective of a large industry (Shell) and deals with *practical criteria* for choosing a neural network solution illustrated by an application for an industrial classification task. The next paper by Braspenning discusses the relationship between NNs and Artificial Intelligence (with its strong emphasis on symbolic processing) and focuses on a *high-level map* of the many types of neural networks and their dynamics, thereby sketching a landscape wherein all treated architectures may be placed. Finally, the contribution by Hudson and Postma addresses the topic of *choosing and using a neural net*, providing suitable criteria in the context of different types of problems, outlining a general categorization of neural network architectures, and finally summing up the considerations that may matter in making a choice.

We would like to finish with a somewhat cautious remark. Although this book is critical at some points about the appropriateness or usefulness of neural network technology, it includes among its purposes that of saving this technology from the sometimes inordinate claims that its enthusiasts are making for it. In a sense, the neural network hype is over! Dressed in more modest but palpable working clothes, neural network technology may yet become a reasonably valuable collection of tools for addressing practical problems.

Finally, we wish to acknowledge those who assisted in making this volume possible. We thank all participants of the NN School and all contributors to this volume. We are grateful to P. Schoo for the computer assistance at the NN School and to J.J.M. Derks and E.J. Pesch for their assistance at various stages of this project. We are especially indebted to Mrs. M. Verheij and Mrs. M. Haenen for preparing this document in LaTeX.

P.J. Braspenning F. Thuijsman A.J.M.M. Weijters

Contents