PERSPECTIVES IN NEURAL COMPUTING

Amanda J.C. Sharkey (Ed.)

Combining Artificial Neural Nets

Ensemble and Modular Multi-Net Systems

Springer

Amanda J.C. Sharkey (Ed.)

Combining Artificial Neural Nets

Ensemble and Modular Multi-Net Systems

With 57 Figures

Springer

Amanda J.C. Sharkey
Department of Computer Science, University of Sheffield, Regent Court,
211 Portobello Street, Sheffield, S1 4DP, UK

Series Editor
J.G. Taylor, BA, BSc, MA, PhD, FlnstP
Centre for Neural Networks, Department of Mathematics, King's College,
Strand, London WC2R 2LS, UK

ISBN 1-85233-004-X Springer-Verlag London Berlin Heidelberg

British Library Cataloguing in Publication Data
Combining artificial neural nets : ensemble and modular
multi-net systems. - (Perspectives in neural computing)
1. Neural networks (Computer science)
I. Sharkey, Amanda J. C.
006.3'2
ISBN 185233004X

Library of Congress Cataloging-in-Publication Data
Combining artificial neural nets : ensemble and modular multi-net
systems. / Amanda J.C. Sharkey, ed.
p. cm. — (Perspectives in neural computing)
Includes bibliographical references.
ISBN 1-85233-004-X (pbk. : alk. paper)
1. Neural networks (Computer science) I. Sharkey, Amanda J. C.,
1957- . II. Series.
QA76.87.C663 1999 98-38717
006.3'2 — dc21 CIP

Typesetting: Camera ready by editor
Printed and bound at the Athenæum Press Ltd., Gateshead, Tyne & Wear
34/3830-543210 Printed on acid-free paper

Perspectives in Neural Computing

Springer
London
Berlin
Heidelberg
New York
Barcelona
Hong Kong
Milan
Paris
Santa Clara
Singapore
Tokyo

Also in this series:

Maria Marinaro and Roberto Tagliaferri (Eds)
Neural Nets - WIRN VIETRI-96
3-540-76099-7

Adrian Shepherd
Second -Order Methods for Neural Networks
3-540-76100-4

Jason Kingdon
Intelligent Systems and Financial Forecasting
3-540-76098-9

Dimitris C. Dracopoulos
Evolutionary Learning Algorithms for Neural Adaptive Control
3-540-76161-6

John A. Bullinaria, David W. Glasspool and George Houghton (Eds)
4th Neural Computation and Psychology Workshop, London,
9-11 April 1997: Connectionist Representations
3-540-76208-6

Maria Marinaro and Roberto Tagliaferri (Eds)
Neural Nets - WIRN VIETRI-97
3-540-76157-8

L.J. Landau and J.G. Taylor (Eds)
Concepts for Neural Networks: A Survey
3-540-76163-2

Gustavo Deco and Dragan Obradovic
An Information-Theoretic Approach to Neural Computing
0-387-94666-7

Thomas Lindblad and Jason M. Kinser
Image Processing using Pulse-Coupled Neural Networks
3-540-76264-7

L. Niklasson, M. Bodén and T. Ziemke (Eds)
ICANN 98
3-540-76263-9

Maria Marinaro and Roberto Tagliaferri (Eds)
Neural Nets - WIRN VIETRI-98
1-85233-051-1

To Autumn and Summer, my lovely daughters.

Preface

The past decade could be seen as the heyday of neurocomputing: in which the capabilities of monolithic nets have been well explored and exploited. The question then is where do we go from here? A logical next step is to examine the potential offered by combinations of artificial neural nets, and it is that step that the chapters in this volume represent.

Intuitively, it makes sense to look at combining ANNs. Clearly complex biological systems and brains rely on modularity. Similarly the principles of modularity, and of reliability through redundancy, can be found in many disparate areas, from the idea of decision by jury, through to hardware redundancy in aeroplanes, and the advantages of modular design and reuse advocated by object-oriented programmers. And it is not surprising to find that the same principles can be usefully applied in the field of neurocomputing as well, although finding the best way of adapting them is a subject of on-going research.

As reflected in the title of this volume, it is possible to make a distinction between two main modes of combining artificial neural nets; ensemble and modular. Under an ensemble approach, several solutions to the same task, or task component, are combined to yield a more reliable estimate. Under a modular approach, particular aspects of a task are dealt with by specialist components before being recombined to form a complete solution. Although their operation differs, both modes can be shown to result in improved performance, and both are represented here. Taken as a whole, the chapters in this volume provide evidence of the advantages of combining nets (by either means). They also explore different methods for creating and combining nets, and provide explanations for their relative effectiveness. This book provides a comprehensive picture of the current state of the art in the new domain of combining Artificial Neural Nets to form multi-net systems. The focus of the book is on combining ANNs, but the methods and results have implications and relevance to the wider machine learning community.

Although they have been revised and updated, versions of some of the chapters (Chapters 1,4,5,6,7,9,10) in the book appeared in two Special Issues of the journal *Connection Science*; Connection Science (1996) 8, 3/4 and Connection Science (1997), 9, 1. *Connection Science* is published by Carfax, P.O. Box 25, Abingdon, Oxfordshire, OX14 3UE.

List of Contributors

Leo Breiman
Statistics Department
University of California
Berkeley, CA 94720, USA

Harris Drucker
Monmouth University,
West Long Branch
NJ, USA

Ione Fine
Center for Visual Science
University of Rochester
Rochester
N.Y. 14627, USA

Cesare Furlanello
ITC-irst, I-38050 Povo (Trento), Italy

Joydeep Ghosh
Dept. of Electrical and Computer Engr.
The University of Texas
Austin, TX 78712-1084,USA

Diego Giuliani
ITC-irst, I-38050 Povo (Trento), Italy

Sherif Hashem
Department of Engineering Mathematics and Physics
Faculty of Engineering
Cairo University
Giza, Egypt

Nathan Intrator
School of Mathematical Sciences
Sackler Faculty of Exact Sciences
Tel-Aviv University
Ramat Aviv 69978, Israel
and
Institute of Brain and Neural Systems
Box 1843,
Brown University,
Providence, RI 02912, USA.

Robert A. Jacobs
Department of Brain and Cognitive Sciences
University of Rochester
Rochester
New York 14627, USA

Stephen P. Luttrell
Defence Evaluation and Research Agency
St Andrews Rd, Malvern
Worcs, WR14 4NL, U.K.

Stefano Merler
ITC-irst, I-38050 Povo (Trento), Italy

David W. Opitz
Computer Science Department
University of Montana
416 Social Science Building
Missoula, MT 591812, USA

Yuval Raviv
School of Mathematical Sciences
Sackler Faculty of Exact Sciences
Tel-Aviv University
Ramat Aviv 69978, Israel

Amanda J.C. Sharkey
Department of Computer Science
University of Sheffield
Sheffield, UK

Jude W. Shavlik
Computer Sciences Department
University of Wisconsin
1210 W. Dayton St
Madison, WI 53706, USA

Martin A. Tanner
Department of Statistics
Northwestern University, USA.

Edmondo Trentin
ITC-irst, I-38050 Povo (Trento), Italy

Kagan Tumer
NASA Ames Research Center
Caelum Research
MS 269-2, Moffett Field, CA, 94035-1000,USA

Table of Contents

1. Multi-Net Systems

Summary.

This chapter provides an introduction to the main methods and issues in the combination of Artificial Neural Nets. A distinction is made between ensemble and modular modes of combination, and the two are then treated separately. The reasons for ensemble combination are considered, and an account is provided of the main methods for creating and combining ANNs in ensembles. This account is accompanied by a discussion of the relative effectiveness of these methods, in which the concepts of *diversity* and *selection* are explained. The review of modular combination outlines the main methods of creating and combining modules, depending on whether the relationship between the modules is co-operative, competitive, sequential or supervisory. An overivew of the chapters in the book forms the conclusion section.

The honeymoon is officially over, and neural computing has moved beyond simple demonstrations to more significant applications. There is a growing realisation that such applications can be facilitated by the development of multi-net systems. Multi-net systems can result in solutions to tasks which either cannot be solved by a single net, or which can be more effectively solved by a system of modular neural net components. Similarly, better performance can be achieved when ANNs, as unstable predictors, are redundantly combined. Arguably, there are few neural net applications accomplished by means of a single net where better performance could not be achieved if this single net were replaced by a multi-net system. As well as performance improvement, there are other advantages to decomposing a task into modular components. For example, a modular system can be easier to understand and to modify. And modularity is almost necessarily implicated in any brain or biological modelling.

It seems likely that multi-net systems will be an important component of future research in neural computing. There are a number of areas from which inspiration and guidance about the construction of such systems can be gained. Clearly we can expect a major contribution from statisticians, and from the wider machine learning community, in terms, for instance, of explanations of the relative effectiveness of different methods for creating and combining ensemble members. Although the focus of concern here is on the combining of artificial neural nets in particular, research on combining other kinds of unstable predictors (e.g. decision trees, see Breiman, Chapter 2, Drucker, Chapter 3) is also relevant, and there is no reason why the members of an ensemble should not consist of a variety of predictors. Insights about combining could potentially be gained from consideration of other areas as well, such as the modelling of biological systems which also make use of redundant and modular elements. And the concept of reliability through redundancy is one that is familiar in a number of different areas, such as software engineering (see Eckhardt [1], for example).

The aim of this chapter is to provide a review of the main methods that have been proposed for combining artificial neural net modules and ensembles, and to examine the principal motivations for creating multi-net systems. However, we shall first turn our attention to a consideration of the distinction between ensembles and modules, and of the ways in which they can be combined to form multi-net systems.

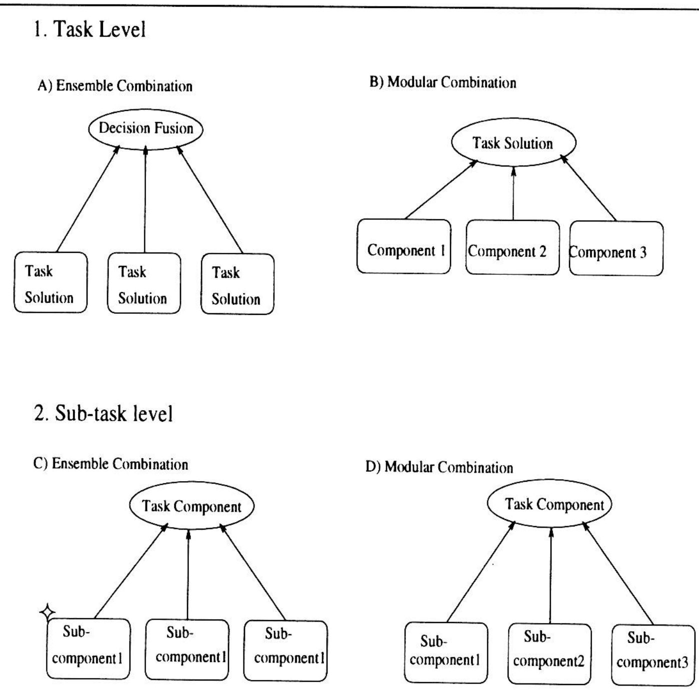

Fig. 1.1. Ensemble and modular multi-net systems, at task and sub-task levels

1.0.1 Different Forms of Multi-Net System

It is useful to make a distinction between ensemble and modular combinations of artificial neural nets [2]. The term 'ensemble' is the one commonly used for combining a set of redundant nets (e.g. [3]), although the term 'committee' [4] or 'committee machine' has also been used for the same purpose. In an *ensemble* combination, the component nets are redundant in that they each provide a solution to the same task, or task component, even though this solution might be obtained by different means. By contrast, under a *modular* approach, the task or problem is decomposed into a number of subtasks, and the complete task solution requires the contribution of all of the several modules (although individual inputs may be dealt with by only one of the modules). Both ensemble and modular combinations can exist at either a task, or a sub-task level, as shown in Figure 1.1.

- Task level: An ensemble could consist of a number of different solutions to an entire task, or problem (Figure 1.1a). Similarly, a task solution might be constructed from a combination of a number of decomposed modules (Figure 1.1b).
- Sub-task level. When a task or an application is decomposed into component modules, each modular component could itself consist of an ensemble of nets, each of which provided a solution for that same modular component (Figure 1.1c). Alternatively, each module could be further subdivided into yet more specialist modules (Figure 1.1d).

At both levels in Figure 1.1, the distinction between an ensemble or modular combination depends on the presence or absence of redundancy; note the redundant components (several versions of Subcomponent 1) in the ensemble sub-task example, (Figure 1.1c) as compared to the lack of redundancy (Subcomponents 1, 2 and 3) in the modular sub-task example (Figure 1.1d). It should be noted that the modular examples (Figure 1.1b and d) at both levels could either result from the decomposition of a task into smaller components, or could represent 'bottom-up' fusion of information from distinct sensors, which provides a link to the quite considerable literature on sensor fusion [see [5] for a review]. Here, rather than decomposing in order to simplify the task, the modular structure can arise as a consequence of the available inputs.

Ensemble and modular combinations should not be thought of as mutually exclusive. It should be noted that Figure 1.1 is designed to show building blocks from which a multi-net system could be constructed: an actual multi-net system could consist of a mixture of ensemble and modular combinations at different levels. As an illustration, Figure 1.2 shows a hypothetical multi-net system which consists of both ensemble and modular components. At the top level, the system consists of an ensemble combination of three task solutions. At the sub-task level however, one of the task solutions is arrived at as the result of a modular combination of distinct components. The three task solutions are produced in different ways. The first is computed on the basis of data from one of three sensors. The second is computed on the basis of a cooperative combination of the output of three sensors. And the third is assembled from the modular combination of three subcomponents, each of which relies on input from a single sensor. Although this figure assumes the