Neural Network Learning: Theoretical Foundations

This book describes recent theoretical advances in the study of artificial neural networks. It explores probabilistic models of supervised learning problems, and addresses the key statistical and computational questions. Research on pattern classification with binary-output networks is surveyed, including a discussion of the relevance of the Vapnik-Chervonenkis dimension. Estimates of this dimension are calculated for several neural network models. A model of classification by real-output networks is developed, and the usefulness of classification with a large margin is demonstrated. The authors explain the role of scale-sensitive versions of the Vapnik-Chervonenkis dimension in large margin classification, and in real estimation. They also discuss the computational complexity of neural network learning, describing a variety of hardness results, and outlining two efficient constructive learning algorithms. The book is self-contained and is intended to be accessible to researchers and graduate students in computer science, engineering, and mathematics.

Martin Anthony is Reader in Mathematics and Executive Director of the Centre for Discrete and Applicable Mathematics at the London School of Economics and Political Science.

Peter Bartlett is a Senior Fellow st the Research School of Information Sciences and Engineering at the Australian National University.

Neural Network Learning: Theoretical Foundations

Martin Anthony and Peter L. Bartlett

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521118620

First published 1999
Reprinted 2001, 2002
This digitally printed version 2009

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data
Anthony, Martin.
Learning in neural networks : theoretical foundations /
Martin Anthony and Peter L. Bartlett.
p. cm.
Includes bibliographical references.
ISBN 0 521 57353 X (hardcover)
1. Neural networks (Computer science). I. Bartlett, Peter L.,
1966– . II. Title.
QA76.87.A58 1999
006.3′2–dc21 98–53260 CIP

ISBN 978-0-521-57353-5 hardback
ISBN 978-0-521-11862-0 paperback

To Colleen, Selena and James.

Contents

Preface

Results from computational learning theory are important in many aspects of machine learning practice. Understanding the behaviour of systems that learn to solve information processing problems (like pattern recognition and prediction) is crucial for the design of effective systems. In recent years, ideas and techniques in computational learning theory have matured to the point where theoretical advances are now contributing to machine learning applications, both through increased understanding and through the development of new practical algorithms.

In this book, we concentrate on statistical and computational questions associated with the use of rich function classes, such as artificial neural networks, for pattern recognition and prediction problems. These issues are of fundamental importance in machine learning, and we have seen several significant advances in this area in the last decade. The book focuses on three specific models of learning, although the techniques, results, and intuitions we obtain from studying these formal models carry over to many other situations.

The book is aimed at researchers and graduate students in computer science, engineering, and mathematics. The reader is assumed to have some familiarity with analysis, probability, calculus, and linear algebra, to the level of an early undergraduate course. We remind the reader of most definitions, so it should suffice just to have met the concepts before.

Most chapters have a 'Remarks' section near the end, containing material that is somewhat tangential to the main flow of the text. All chapters finish with a 'Bibliographical Notes' section giving pointers to the literature, both for the material in the chapter and related results. However these sections are not exhaustive.

It is a pleasure to thank many colleagues and friends for their contri-

butions to this book. Thanks, in particular, to Phil Long for carefully and thoroughly reading the book, and making many helpful suggestions, and to Ron Meir for making many thoughtful comments on large sections of the book. Jon Baxter considerably improved the results in Chapter 5, and made several useful suggestions that improved the presentation of topics in Chapter 7. Gábor Lugosi suggested significant improvements to the results in Chapter 4. Thanks also to James Ashton, Shai Ben-David, Graham Brightwell, Mostefa Golea, Ying Guo, Ralf Herbrich, Wee Sun Lee, Frederic Maire, Shie Mannor, Llew Mason, Michael Schmitt and Ben Veal for comments, corrections, and suggestions. It is also a pleasure to thank the many collaborators and colleagues who have influenced the way we think about the topics covered in this book: Andrew Barron, Jon Baxter, Shai Ben-David, Norman Biggs, Soura Dasgupta, Tom Downs, Paul Fischer, Marcus Frean, Yoav Freund, Mostefa Golea, Dave Helmbold, Klaus Höffgen, Adam Kowalczyk, Sanjeev Kulkarni, Wee Sun Lee, Tamás Linder, Phil Long, David Lovell, Gábor Lugosi, Wolfgang Maass, Llew Mason, Ron Meir, Eli Posner, Rob Schapire, Bernhard Schölkopf, John Shawe-Taylor, Alex Smola, and Bob Williamson. We also thank Roger Astley of Cambridge University Press for his support and for his efficient handling of this project.

Parts of the book were written while Martin Anthony was visiting the Australian National University, supported by the Royal Society and the Australian Telecommunications and Electronics Research Board, and while Peter Bartlett was visiting the London School of Economics. Martin Anthony's research has also been supported by the European Union (through the 'Neurocolt' and 'Neurocolt 2' ESPRIT projects) and the Engineering and Physical Sciences Research Council. Peter Bartlett's research has been supported by the Australian Research Council and the Department of Industry, Science and Tourism. We are grateful to these funding bodies and to our respective institutions for providing the opportunities for us to work on this book.

We thank our families, particularly Colleen and Selena, for their help, encouragement and tolerance over the years.

Martin Anthony and Peter Bartlett
London and Canberra
March 1999.